Problem Chosen	2023 MCM/ICM Summary Sheet	Team Control Number
A		2308434

Numerous Species in ONE Model: Effects of Biodiversity on Environmental Resistance

Summary

Recent years have witnessed occurrences of extreme weather events in nature, containing droughts, which vary in frequency and intensity. Plants, reacting to stresses in various ways, show **greater environmental resistance as the advance of biodiversity**. Consequently, it is significant for us to analyze this phenomenon.

Firstly, to actualize our study, we collect data from various databases and choose the location and specimens for simulation. **The Shannon Wiener index and Simpson's Dominance index** are used for evaluation of the development of a community and its biodiversity.

Secondly, after quantifying the effects of temperature and precipitation on the community, we construct **the Ideal Plant Community Growth Model** Φ by using **the correlation matrix from Gause's team** to measure the interactions between species, where we analyze the **diagonal** impacts on individual species and the **off-diagonal** relationships (Competitive and Symbiotic) to determine the value of the elements in Φ. We also predict the development of the community through **Quasi Markov Chain**.

Thirdly, choosing heat waves and droughts to examine the resistance of the community, we find out that it can only maintain the utmost of its species' survival with more than **4 species**. Moreover, we compare communities consisting of symbiotic and competitive species with the same number of species, which attaches great importance to interactions. As a result, **the rise of biodiversity can strengthen the environmental resistance of a community**.

Ultimately, as for weather in the future, we choose the **ARIMA model** and **Ordinary Least Squares** and perfectly predict the regular weather over the next century, with **RMSE lower than 10%**. Then we use the predicted data to simulate the community reaction to varying temperatures and precipitation, from which we conclude that a community with numerous species can **protect vulnerable ones from extinction**. After that, we discuss the influences of pollution and habitat reduction and find out that these factors prone to harm biodiversity dramatically. As a matter of fact, we propose necessary measures to avoid them and analyze their impacts on the larger environment.

Keywords: ELS (Exponential Least Squares), Gause's model, Shannon Wiener Index, Quasi Markov chain, ARIMA model, Biodiversity

Contents

1 Introduction

1.1 Problem Background

Plants have developed various adaptive mechanisms in response to stresses, as shown in Fig.1. In recent years, there have been numerous occurrences of extreme weather events in the natural world, including droughts, which exhibit significant variations in frequency and intensity. As such, it has been observed that a greater diversity of plant species within a community may confer greater resistance to drought, particularly over repeated cycles of exposure spanning multiple generations (shown in Fig.2) Accordingly, it is pertinent to explore the minimum number of plant species required to let a plant community benefit, the impact of increased species diversity, as well as the implications of these factors for long-term survival.

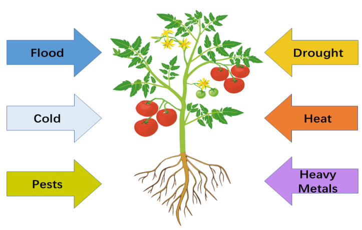


Figure 1: Plant under stress

Figure 2: Drought plant community [1]

1.2 Restatement of the Problem

Based on the background information, the main tasks of this paper are as follows:

• Construct a mathematical model for the prediction of changes in a plant community over time when various irregular weather cycles strike, including extreme drought and flood disasters. We should consider interactions between diverse species exposed to cycles of drought.

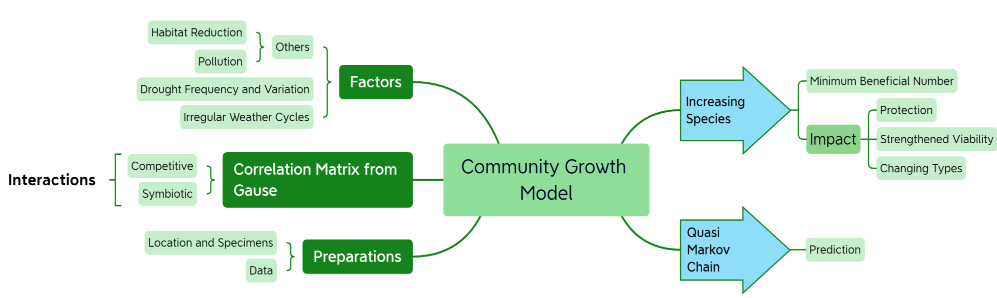
• Find out the indispensable number of plant species to maintain the survival of most of the species in the community and things taking place as the number increases. We need to discuss the influence of types of species on the results of prediction as well.

• Explore what happens when the occurrence of future droughts in weather cycles strikes more frequently and varies more widely. We are also supposed to compare the impact of the number of species on the whole population in conditions of less frequent droughts with the original results.

• Take other factors into account, such as pollution and habitat shrinking.

• Make a list of necessary measures to let a plant community survive for a long time and analyze their effects in a more extensive environment.

1.3 Our Work



2 Global Assumptions and Justifications

- The community is in a weather-evenly-conditioned circle area, the radius of which is R.

  ⇒**Justification:** The shape of the community counts little, so we suppose an ideal area so as to simplify the problem.

- The community benefit if most of the species can survive under irregular weather conditions.
  ⇒**Justification:** We can give a quick proof by contradiction: suppose the death of some species in the community will benefit the survival of all, then the size of the community will diminish alongside the iteration, so in a relatively long time, no species will survive, where we reach a contradiction.

- In our model, we consider that pollution and habitat reduction are mainly caused by human activities.
  ⇒**Justification:** Analysing the data from Gui $et\ al.$ [2], we made our assumption accordingly. And this assumption corresponds to our intuition.

3 Model Preparation

3.1 Notations

Table 1: Variables used in this section

Symbol	Definition
α	Community Development Index
T	Temperature
P	Precipitation
F(T,P)	Parameter of Environmental influence
φ	Parameter of Interactions in community
H	Shannon Wiener index
D	Simpson's dominance index

3.2 Data Compilation

The data we used mainly contains climate characteristics, plant features, and historical extreme weather data. The data sources are summarized in TBL.2.

3.3 Location and specimen

- Location of the Community

Table 2: Data Source Collation

Database Names	Database Websites	Data Type
CLIMATES TO TRAVEL	https://www.climatestotravel.com/climate/united-states/kansas-city	Geography
USDA	https://plants.usda.gov/home	Biology
National Climate Assessment	https://nca2014.globalchange.gov/	Geography
NOAA	https://www.noaa.gov/	Geography

To actualize our work, we utilize the climate characteristics of the grassland in the Great Plains, where the center of the assumed circle area (shown as the circle in Fig. 3) is at Kansas City (Missouri).

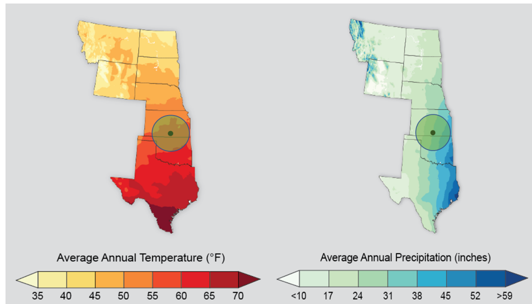
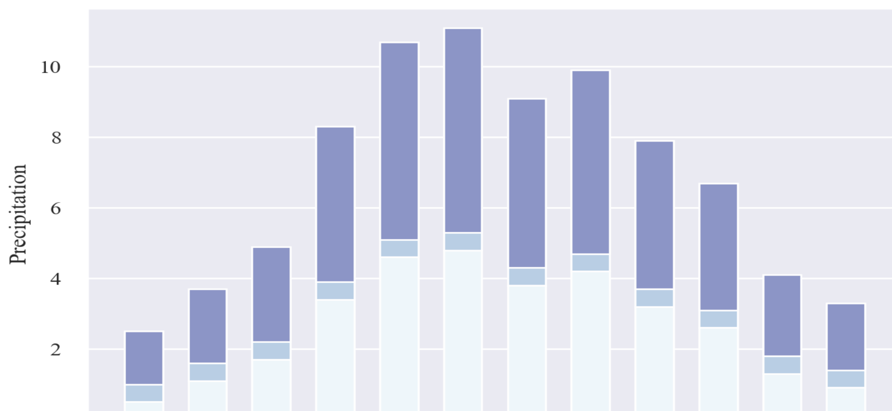


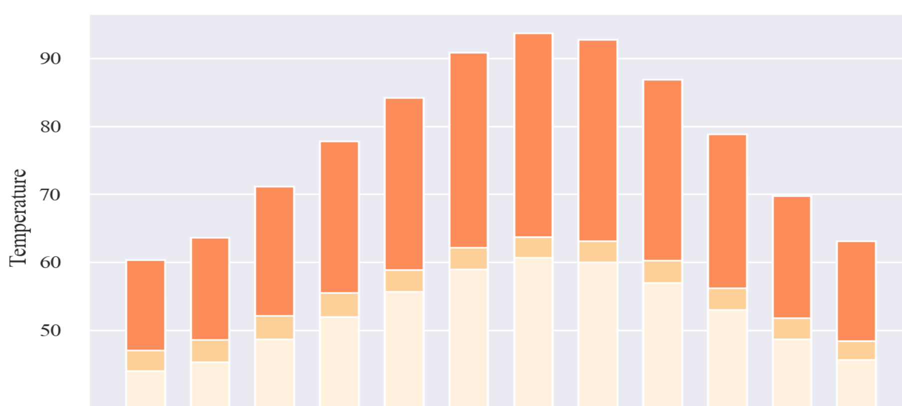
Figure 3: Average annual temperature and precipitation of the Great Plains [3]

Then we get the annual data of the temperature and precipitation of the Great Plains from the site *CLIMATES TO TRAVEL* (listed above), as is shown in Fig. 4 and Fig. 5.



	January	February	March	April	May	June	July	August	September	October	November	December
max	1.5	2.1	2.7	4.4	5.6	5.8	4.8	5.2	4.2	3.6	2.3	1.9
mean	1.0	1.6	2.2	3.9	5.1	5.3	4.3	4.7	3.7	3.1	1.8	1.4
min	0.5	1.1	1.7	3.4	4.6	4.8	3.8	4.2	3.2	2.6	1.3	0.9

Figure 4: Annual precipitation of the Great Plains



	January	February	March	April	May	June	July	August	September	October	November	December
max	40.0	45.0	57.0	67.0	76.0	86.0	90.0	89.0	80.0	68.0	54.0	44.0
mean	31.1	35.8	46.4	56.5	66.7	76.5	81.1	79.2	70.7	58.5	45.4	35.3
min	22.0	26.0	36.0	46.0	57.0	67.0	72.0	70.0	61.0	49.0	36.0	27.0

Figure 5: Annual temperature of the Great Plains

- Species We Choose for Simulation

Then we choose the six most common plant species on the Great Plains of North America as simulation specimens, which are numbered sequentially from S_1 to S_6, to express the interactions of multiple species. Their names, numbers, and characteristics can be found in TBL.3.

Table 3: Names, numbers, and characteristics of the six species

Name	Number	Height	Growth Characteristics	Higher Drought Resistance	Temperature Preferences
Big Bluestem	1	tall	grows in dense stands	✘	warm
Little Bluestem	2	short	often grows in clumps	✘	warm
Indiangrass	3	tall	has tall, showy seed heads	✘	warm
Western Wheatgrass	4	short	grows in bunches	✘	cool
Prairie Junegrass	5	short	grows on rocky slopes	✔	cool
Blue Grama	6	short	grows in dense clumps	✔	warm

As shown in the table, we numbered them from 1 to 6 for convenience. In the following part, we will only refer to S_i, instead of its scientific name.

4 Model I: Ideal Plant Community Growth Model

4.1 Metric for Biodiversity of a Community

To evaluate the development of a community and its biodiversity, we use the following evaluation method.

Define α_i as the number of individuals of S_i, and A_i as the environmental carrying capacity of S_i. Limited by necessary resources including water and space, the upper bound of growth for population development within a limited environment should be a constant, which is to say, $A_i = \|\alpha\|_\infty = sup(\alpha_i)$.

Assuming that the community contains n species, we define $\alpha_{relative} = \frac{\sum_{i=1}^{n} \frac{\alpha_i}{A_i}}{n}$ as the integrated community development index denoted as α_{re}.

Since the difference in the number of species in a grassland community is not huge, we can also use $\alpha_{absolute} = \sqrt{\sum_{i=1}^{n} \alpha_i^2}$ to approximate the comprehensive development of the community, which is dennoted as α_{ab}.

In addition, to evaluate the biodiversity and species evenness of the community, we introduce two concepts.

• Shannon Wiener index:

$$H = -\sum_{i=1}^{n} P_i \ln(P_i)$$

where $P_i = \frac{\alpha_i}{\sum_{j=1}^{n} \alpha_j}$ measures the proportion of population i in the whole community.

The Shannon-Wiener index defines the "information entropy" of the community, from the perspective of information theory . The larger the Shannon-Wiener index is, the greater the heterogeneity of the community and thus the information disorder and the uncertainty, and the greater the species richness. Shannon-Wiener index is 0 indicates that there is only one species in the community.

• Simpson's dominance index:

$$D = \sum_{i=1}^{n} P_i^2$$

It is defined over a probabilistic view. The larger the Simpson's dominance index, the more homogeneous the community and the more concentrated the species distribution.

4.2 Quantification of Environmental Impacts

In order to show the impact of the exterior environment, we construct $F(T, P)$ as a function of temperature and precipitation. We map the data from a high dimensional space to the R-One uniform space, the procedure of which can be illustrated in Fig.6:

We denote m and n as the sample frequency (12 months in our case) and the number of specimens (6 plants). The subsets of T and P are then mapped from the feature space to the evaluation space (where exists $\tilde{T}$ and $\tilde{P}$) via the Normalized Gaussian function, and σ_1^2 , $\sigma_2^2$are the standard deviation respectively. Ultimately, we are required to construct a mapping function $F\langle *, * \rangle$, which should be applied to the following constraints:

$$\begin{cases} \lim_{* \to 0} F\langle *, * \rangle = 1 \\ \lim_{* \to 1} F\langle *, * \rangle = 0 \\ H_F \succeq 0 \end{cases}$$

The former two equations represent the boundary conditions, while the latter stands for the semi-definite positive restriction of the function $F\langle *, * \rangle$, and we calculate the seconde dirivitive Hessian matrix of the function. Weighing the pros and cons, we construct the following function :

$$\begin{cases} F(T, P) = e^{a - b\sqrt{\tilde{T} \cdot \tilde{P}}} + c \\ \tilde{T} = e^{-\frac{\sum (T - T_2)^2}{2\sigma_2^2}} \\ \tilde{P} = e^{-\frac{\sum (P - P_2)^2}{2\sigma_2^2}} \end{cases}$$

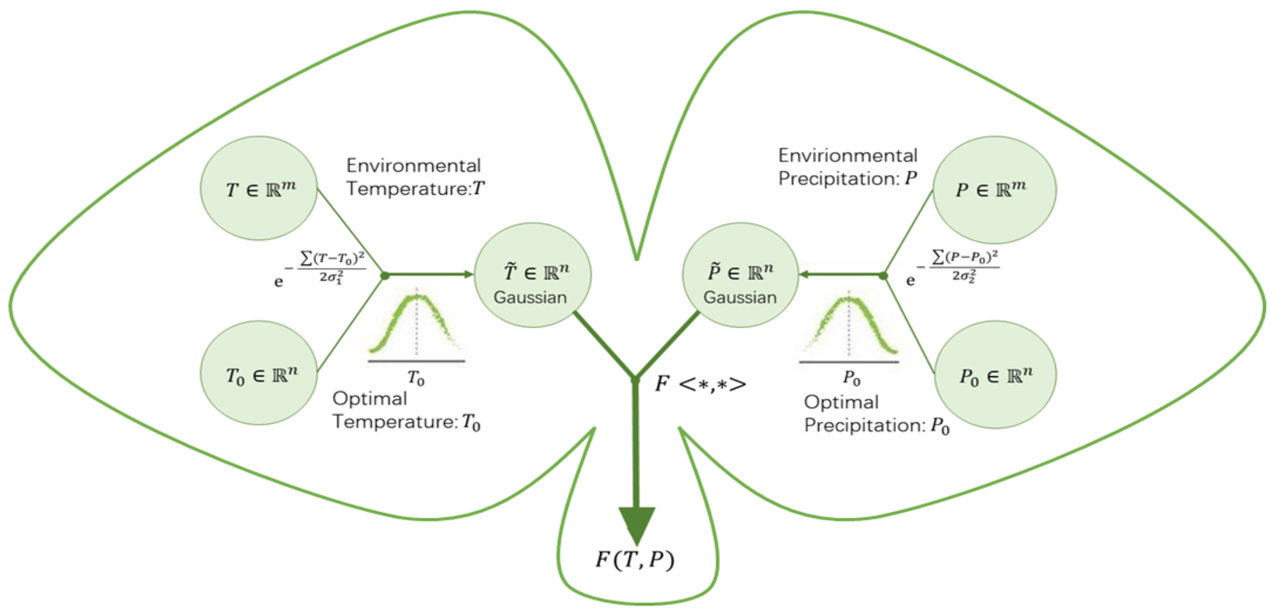


Figure 6: Derivation process of $F(T, P)$

And the problem reduces to an argmin algorithm (where x_i stands for $\sqrt{\tilde{T}_i \cdot \tilde{P}_i}$):

$$\underset{a,b,c}{\operatorname{argmin}} \sum_i \left(e^{a-bx_i} + c - F\left(x_i\right)\right)^2$$

The ELS (Exponential Least Square) method is applied to fit the data. Let us consider three data points indexed by $i,\ j,\ k$:

$$\frac{F_{x_i} - F_{x_j}}{F_{x_i} - F_{x_k}} = \frac{e^{a-bx_i} - e^{a-bx_j}}{e^{a-bx_i} - e^{a-bx_k}} = \frac{1 - e^{b(x_j - x_i)}}{1 - e^{b(x_k - x_i)}}$$

Now, select $x_k = \frac{x_i + x_j}{2}$, which leads to

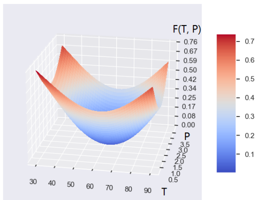
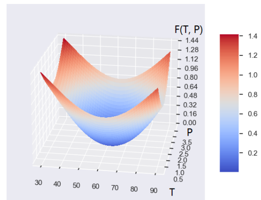
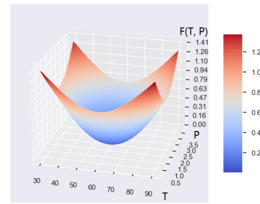
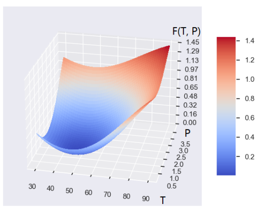
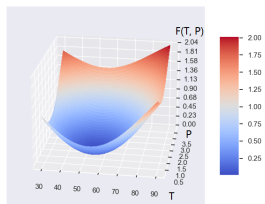
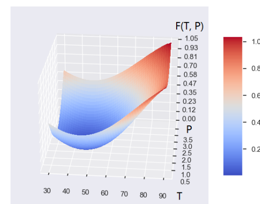
$$F' = \frac{F_{x_k} - F_{x_j}}{F_{x_i} - F_{x_k}} = e^{b\frac{x_j + x_i}{2}} = e^{bx_k}$$

We take the logarithm of the equation and multiply the function vector F' by the pseudoinverse of coefficient matrix (Obviously it has independent columns) and thus we determine the $\hat{b}$:

$$\begin{bmatrix} 0 \\ \hat{b} \end{bmatrix} = \left([\mathbf{1}\ \mathbf{X}]^\top [\mathbf{1}\ \mathbf{X}]\right)^{-1} \cdot [\mathbf{1}\ \mathbf{X}]^\top \cdot [\ln(\boldsymbol{F'})]$$

In this manner, we can calculate the a and c without hassle, and a, b, c are 0.48867, 1.704390, 0.58198 respectively.

Table 4: The variation of $F(T, P)$ for different species

Big Bluestem	Little Bluestem	Indiangrass
		
Western Wheatgrass	**Prairie Junegrass**	**Blue Grama**
		

We input the data into our function. The optimal temperature and precipitation for different species vary in a range:

As we can see in TAB.4, the value of $F(T, P)$ touches 0 when the temperature is near 55 $^\circ F$ and the precipitation is near 2.5 inches (the dark blue part), which indicates the influence of the external environment reaches the minimum, in other words, the optimal environmental conditions locate in the dark blue part.

Conversely, when the temperature approaches 90 $^\circ F$ or 30 $^\circ F$ while the precipitation reaches 0 inch or 4.0 inches, $F(T, P)$ reaches the peak, as in the red part. That signifies the maximum impact of outside, as well as the worst environmental conditions.

As a result, we can plot the variation of $F(T, P)$ over one year via our model in Fig.7:

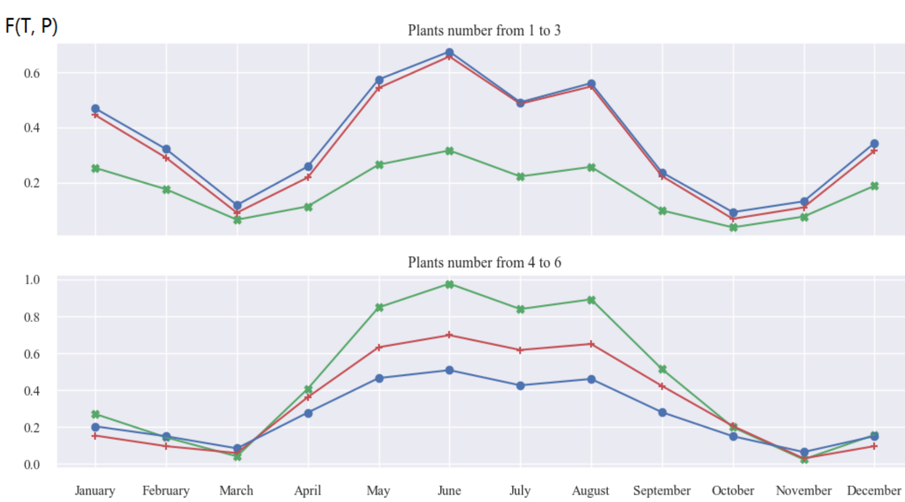


Figure 7: The variation of $F(T, P)$ in one year

4.3 Quantification of Interactions Between Species

While analyzing the growth of a plant community, the interactions between species count, which we are thus supposed to quantify. According to the Gause team's result [4], the correlation matrix can be written as follows:

$$\Phi = (\varphi_{ij})_{n\times n} = \begin{bmatrix} \varphi_{11} & \varphi_{12} & \cdots & \varphi_{n1} \\ \varphi_{21} & \varphi_{22} & \cdots & \varphi_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ \varphi_{n1} & \varphi_{n2} & \cdots & \varphi_{nn} \end{bmatrix}$$

We denote α_i as the biomass of the i^{th} species, which is mainly determined by the population, as is explained in $Collins$ [5]. Then the effect of S_i on S_j can be denoted as $\varphi_{ij}\alpha_i$, where φ_{ij} is a function depending on α indicating the relationship between species. Fig.8 indicates that the elements of matrix Φ vary with time. It is crystal clear that the elements on the diagonal counts more and fluctuate more violently. In the following part, we are to explain the idea of this correlation matrix in detail:

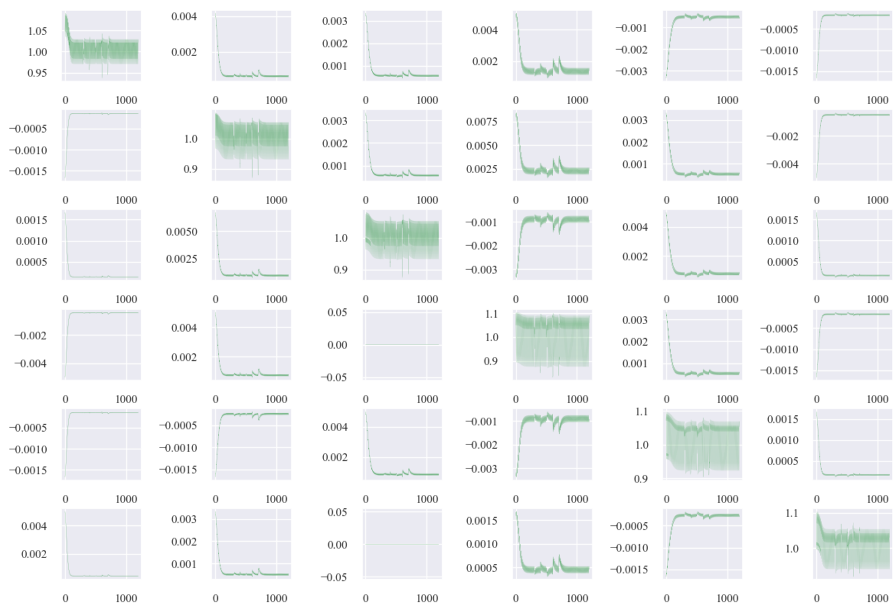


Figure 8: Elements of Φ vary with time

• Diagonal

The constants on the diagonal stand for two parts: the intrinsic growth rate and the impact from the exterior environment. From the research of D. Lobell, K. Nicholas, and C. Field [6], they can be depicted as:

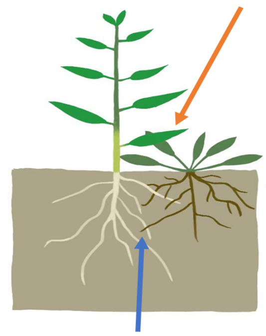
$$\varphi_{ii} = r_i(1 - \frac{\alpha_i}{A_i} - F(T, P))$$

Where A is the carrying capacity in the Logistic curve, and $F(T, P)$ is the function above.

• Off-diagonal

↪ If the two species compete, φ_{ij} is negative because of the following competitions (shown in the figure):

• Competition for inorganic compounds, owing to the alike depths to which the roots of the two species may extend.
• Competition for water, because the water absorption capacity and thus the water reserves of soil in a restricted area are similar.
• Competition for light, as their leaves may overlap and shade each other.



↪ Conversely, if the two species are symbiotic, then φ_{ij} is positive for the following reasons:

* From USDA (listed above) we know that some herbaceous plants are likely to grow under shades or in dense stands, taking Big Bluestem as an example. Consequently, one species of a specific plant density may promote the growth of the other.

* Thanks to the allopathy of plants, a plant can release chemicals that have positive effects on another one, as is indicated in $Collins$ [7].

Then we are able to get the interactions with the whole community. The j^{th} species can be influenced by every species (including itself), and the $ODE's$ (Ordinary Differential Equation) for our prediction model are as follows:

$$\frac{d\alpha_j}{dt} = \sum_{i=1}^{n} \varphi_{ij}\alpha_i$$

We can also set α as an n-dimensional vector and rewrite it in an algebraic form:

$$\frac{d\alpha}{dt} = \Phi \cdot \alpha$$

It is proven to satisfy the Lipshitz condition:

$$\begin{cases} \left\|\Phi \cdot \alpha - \Phi \cdot \alpha'\right\|_2 \leqslant M \left\|\alpha - \alpha'\right\|_2 \\ M = \sup\left\{\max\left\{\varphi_{ij}\right\}, \max\left\{\frac{r_i}{A_i}\right\}\right\} \end{cases}$$

It appears that our model utilizing a convex function has satisfied the Lipschitz condition to a reasonable degree. As a matter of fact, we can use the law of Liouville:

$$W_{(t)} = W_{(t_0)} e^{\int_t^t \operatorname{tr}\left(\Phi^\top \Phi\right) dt}$$

Where $W_{(t)}$ stands for the determinant of the fundamental solution matrix to the $ODE'\ s$ above. In order to solve the $ODE'\ s$ by numerical computation method, we discretize the equation with $Forward\ Finite\ Difference\ method$, from which we are able to indicate:

$$\alpha_{(t=k+1)} - \alpha_{(t=k)} \approx \Phi \cdot \alpha_{(t=k)} \cdot \Delta t$$
$$\alpha_{(t=k+1)} = (\Phi \Delta t + I) \cdot \alpha_{(t=k)}$$

We denote $(\Phi\Delta t + I)$ as Θ, the transition matrix of our $Quasi\ Markov\ chain$.

$$\alpha_{(t=k+1)} = \Theta^{k+1} \cdot \alpha_{(t=0)}$$

In this manner, we can predict the variation of the total biomass of the ideal community by multiplying the transition matrix over and over in a relatively computational-friendly way.

5 Model II: Ideal Plant Community under Irregular Weather Cycles

After analyzing the growth of a community under ideal conditions, we impose extreme weather on the species. In this case, the impact of internal factors such as the number and type of the species should be analyzed.

5.1 Ideal Plant Community under Drought or Torridity

Being exposed to extreme conditions (taking drought for instance), our six-specimen-community demonstrates its relatively high resistance to trauma, and the result is shown below.

The top two panels show how our six specimens behave in the face of varying degrees of heat. The top and the bottom show curves under droughts, while the left and right parts represent mere and intense abnormal weather, respectively.

As Fig.10 demonstrated, due to the differences in the properties of the six plants we studied, they behave differently in the face of adversity. For example,S_6 (the orange curve) usually acts as the dominant species and even grows better under small heat waves. In addition to the decline of species it competes with, the reason contains its excellent endophytic growth rate and characteristics as a typical warm season plant. The same reason could also explain why S_2 and plant S_5 get a small buff when encountering typical dry weather (the blue and red curves in the lower-left graph). On the contrary, take S_4 for instance, which is a typical cool season plant with relatively high water demand, so it underperforms both under severe drought and heat, leading to the fact that S_6, with which it

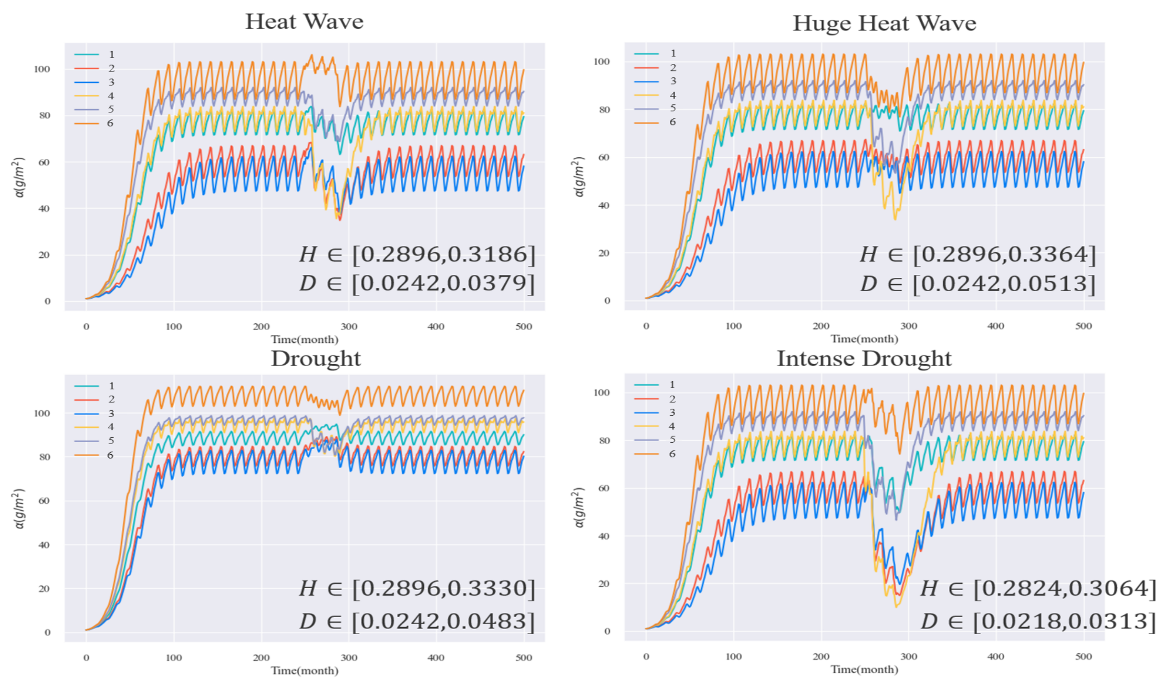


Figure 9: Biomass growth curve under one-time irregular weather cycle

has a symbiotic relationship, although it has excellent drought resistance properties, still suffers more on condition of extreme droughts.

In fact, the characteristics of different species in a real grassland community and their interactions are more complex than those simulated by our model, whereas our model reveals the fact that different species in a diverse and well-structured community react differently to environmental changes, and such differences are propagated throughout the community via its complex interaction network. In the majority of cases, such effects are propagated on a relatively large scale, canceling each other out and attenuating various risks in propagation, thus making the whole community more robust to irregular weather cycles. We will notice this kind of phenomenon in the results of Model 2.

5.2 The Minimum Number of Species to Benefit the Community

We have chosen six species for convenience for the study above, and then we will find out the minimum number of species that can let the community benefit. As written in global assumptions, we reckon that a community benefits when most of its species are able to survive after numerous cycles of irregular weather.

We set the number of species as an independent variable while simulating the value of α. If α_i reaches 0, then S_i has died out. As Fig. 10 illustrates, we impose a drought on the community at the 300^{th} month to examine its drought resistance. After the curve returns stable, only communities with greater than or equal to 4 species ensure the survival of all species, while the rest loses more than one-third of species. As a result, it is notable that the community benefits when there are at least **4 species**.

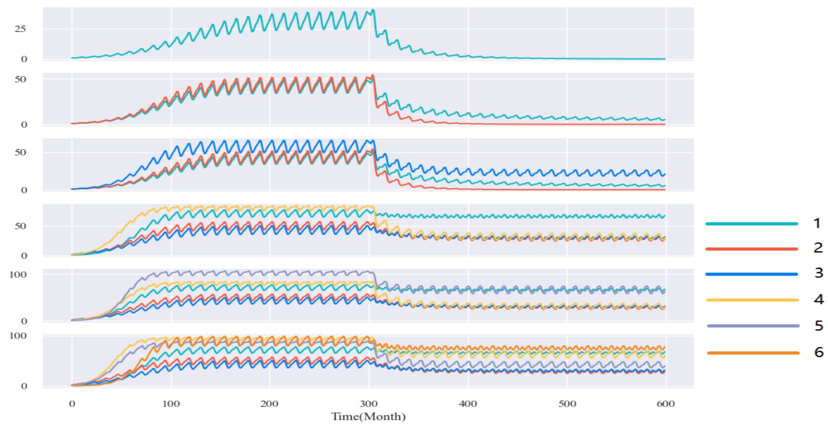


Figure 10: Simulation with different numbers of species

When the number of species is less than or equal to three, there is only one species left in the communities that experienced drought weather, and the communities exist in name only. When the number of species is greater than three, according to the calculation, the various indexes of the community before and after the disaster are shown in the table.

Communities	H before	H after	change ratio	D before	D after	change ratio
C_4	$0,3623$	$0,3028$	-16.42%	$0,0427$	$0,1813$	324.59%
C_5	$0,3336$	$0,3312$	-0.72%	$0,0276$	$0,0895$	224.28%
C_6	$0,3662$	$0,3653$	-0.25%	$0,0086$	$0,0478$	455.81%

As you can see, as the number of species increases, on the one hand, the H index of the community does not change significantly, but the ability of the species diversity of the community to resist disasters is improved; on the other hand, the D index decreases significantly, which means that the community The dominant species are rapidly losing their dominance as the number of species increases, and the community structure is becoming reasonable.

5.3 Impact of Types of Species

The abilities of different species are illustrated in the radar graph below.

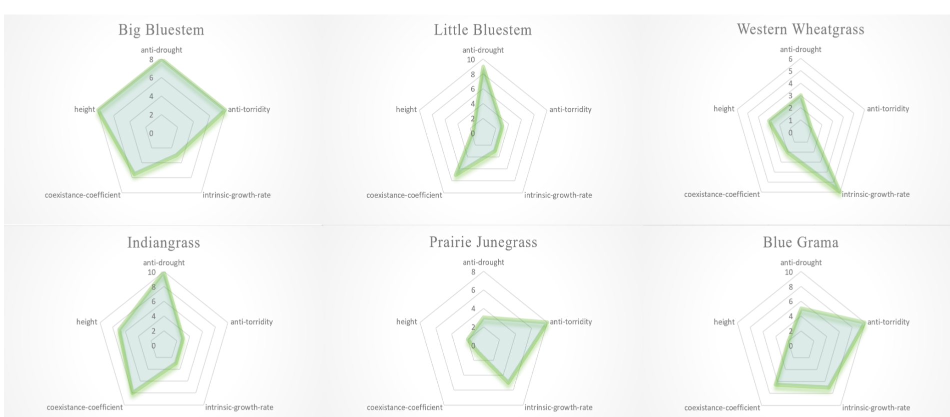


Figure 11: Ability Value of each plant

Where height measures the height of an individual plant, anti-drought shows the ability to resist droughts, anti-torridity is the capacity to resist torridity, intrinsic-growth-rate represents the intrinsic growth rate (r) of an individual plant, and coexistence-coefficient is denoted as the competence of affecting the other species in the same community.

Their interactions can also be illustrated in the Sankey diagram in Fig. 12:

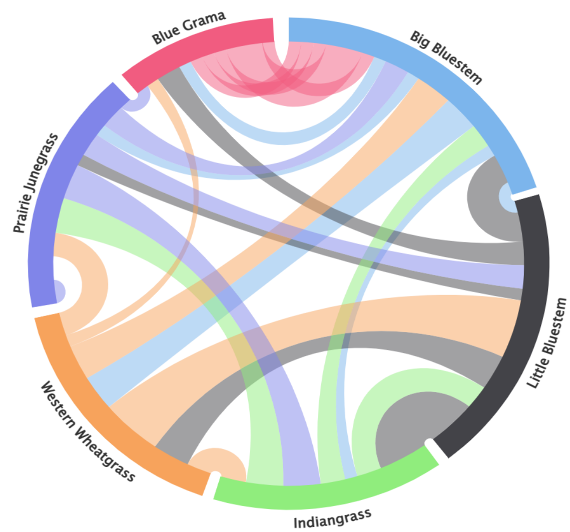


Figure 12: Sankey Diagram: Interaction of different types

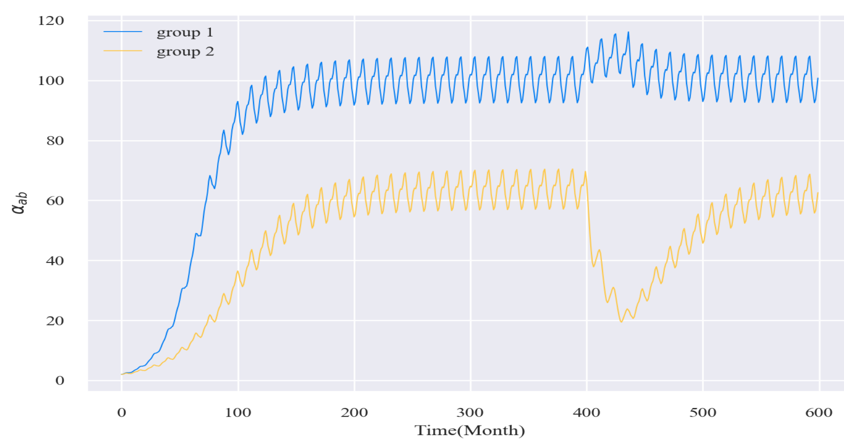


Figure 13: The performance of structurally different communities in the face of drought

Group1 and Group2 in Fig.13 contain four species (plant 1 2 4 6 vs. plant 2 3 4 5), respectively. We assume that at the 300^{th} month - when, obviously, both communities have reached a steady state - a dry spell lasting two years strikes (with an average increase in temperature of 10% and an average decrease in precipitation of 30%), and it is easy to find out that Group2 is far more vulnerable than Group 1. Over the month with the greatest impact, the value of α_{ab} in Group1 is 13.93% higher than that at steady state, while group2 was 69.79% lower. This is due to the fact that most of the plants in Group 1 are warm-season plants and more drought-tolerant. On the contrary, half of the plants in Group 2 are cool-season plants and have relatively higher precipitation demand, so they behave more sensitively in the face of environmental changes.

In addition, plant interactions should be taken into account. In Group 1, plants are mainly symbiotic or do not interfere with each other, but in Group 2 the species compete intensely. This has led to different developmental pictures of the community.

A glimpse of the whole picture shows that different species compositions with the same number of species can also have a great impact on community performance. The reasons for this are, on the one hand, the nature of different species that lead to different overall community properties and, on the other hand, the different interactions between species. Predictably, when the number of species and the size of the community increase, the overall properties of the community are no longer significantly affected by the properties of individual species, while the structure of species interactions within the community becomes the decisive factor.

5.4 Consequences of increasing species

According to our analysis in 5.2, as the number of species increases, the biodiversity of the community is enhanced. As a result, the resistance stability is strengthened, leading to the reinforcement of long-term viability under irregular weather cycles.

In fact, the characteristics of different species in a real grassland community and their interactions are more complex than those simulated by our model, whereas our model reveals the fact that different species in a diverse and well-structured community react differently to environmental changes, and such differences mingle in the community via its complex interaction network. In most cases, such effects take place at a relatively large scale, crippling various risks in propagation, and thus making the whole community more robust to irregular weather cycles. This phenomenon is to be discussed in the following analysis.

6 Analysis of the External Factors

6.1 More Droughts: Irregular in Frequency and Variation

The historical dataset for our study area is denoted as y_t. Specifically, we focus on the annual average temperature change series from 1991 to 2010, which is illustrated together with our prediction result in Fig. 16.

It can be learned that the data vibrates weakly with no clear trend.

We decide to choose the $ARIMA(p, d, 0)$ model to predict the temperature fluctuation, so that we get:

$$\Delta^{(d)} y_t = \sum_{j=1}^{p} \beta_j \Delta^{(d)} y_{t-j} + \epsilon_t$$

Where $\Delta^{(d)}$ performs a d-order difference on the sequence y_t and $\epsilon_t \sim N(0, \varepsilon_t^2)$ is denoted as the residual. Thus, we obtain the first-order difference prediction of the series in t years:

$$E(\Delta^d y_t) = \sum_{j=1}^{p} \hat{\beta}_j E(\Delta^{(d)} y_{t-j})$$

Then, concerning the assumptions of the ARIMA model and empirical experience, the predicted values satisfy the normal distribution $y_t|y_{t-1} \sim N(\mu_t, \sigma^2)$. As a result, considering the stochasticity, the temperature predicted at the moment t can be modified as follows:

$$\hat{y}_t = E(\Delta^{(d)} y_i) + \hat{y}_{t-1} + \epsilon_t$$

In the stepwise prediction process, numerous results may be obtained due to the accumulation of randomness.

Referring to the OLS (Ordinary Least Squares) method in the linear regression model, we exploit linear expression for this prediction model:

$$\begin{bmatrix} \Delta y_p & \Delta y_{p-1} & \dots & \Delta y_1 \\ \Delta y_{p+1} & \Delta y_p & \dots & \Delta y_2 \\ \vdots & \vdots & \ddots & \vdots \\ \Delta y_{240} & y_{239} & \dots & \Delta y_{240-p} \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_p \end{bmatrix} = \begin{bmatrix} \Delta y_{p+1} \\ \Delta y_{p+2} \\ \vdots \\ \Delta y_{p+240} \end{bmatrix}$$

We rewrite it in an algebraic form $(X^T X)\boldsymbol{\beta} = X^T Y$. Consequently, the closed-form solution for the corresponding parameter as follows(Obviously X has independent columns):

$$\hat{\boldsymbol{\beta}} = X^{\dagger} Y = (X^T X)^{-1} X^T Y$$

Therefore, the OLS method can be used to estimate the parameters of the formula by combining the truncated tails of the autocorrelation coefficient (ACF) and the partial correlation coefficient (PACF) of the final differential data (shown in Fig. 14 and Fig. 15) with the trailing case.

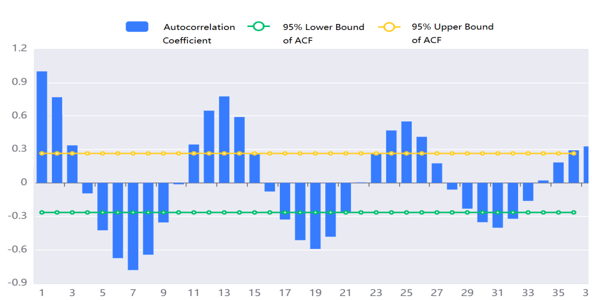


Figure 14: ACF

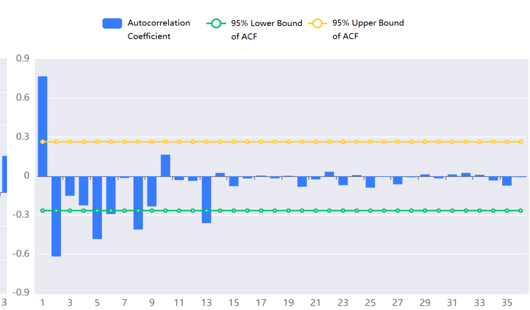


Figure 15: PACF [1]

In the $ARIMA$ model, there exist lagged and differential orders. To ensure the accuracy of our model, we calculate the **RMSE (Root Mean Squared Error)** of our model, which is smaller than the lowest criteria $\epsilon = 10\%$ for high fitting precision. As a matter of fact, we can affirm that we fit the data well.

The value of $RMSE$ is shown as follows:

$$\epsilon = \sqrt{\frac{1}{l}\sum_{i=1}^{l}\left(y_i - \hat{y}_i\right)^2} = 3.5569\% < 10\%$$

The final result is shown in Fig. 16. It can be seen that our model fits the original data well and exactly predicts future situations.

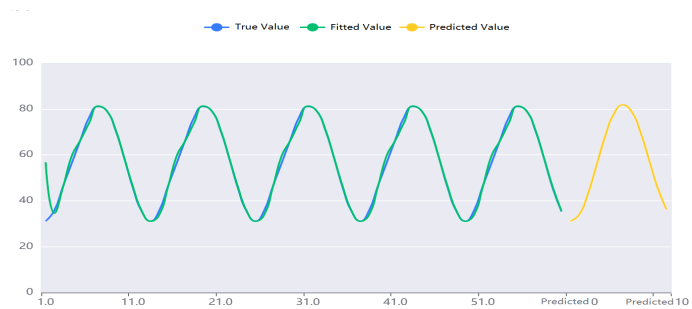


Figure 16: Presentation and prediction of data by our model

By iteratively applying the aforementioned methodology to the precipitation data, the climate data for the studied region can be ascertained for the ensuing century.

It should be noticed that the data serving as the basis for our prediction does not contain information about the occurrence of some unusual climate. It is stable and regular, generating predictions that show the same stability and lacking variability. Indeed, however, it is the random occurrence, especially major abnormal climate phenomena, that influences the studied plant communities the most. Therefore, we examined anomalous weather conditions in this area over the last century (shown in TBL.5).

Table 5: Anomalous weather conditions for the last century

Name of the disaster	**Main Impact of disaster**	**Time of disaster**
Arkansas River Flood	P $\nearrow$	1927
Dust Bowl	T $\nearrow$ P $\searrow$	1930s
Abnormal Hot Weather	T $\nearrow$	1936
Wichita River Flood	P $\nearrow$	1951
Midwest Blizzard	P $\searrow$	1977
Wichita Heat Wave	T $\nearrow$	1980
Scorching Heat in2010s	T $\nearrow$	2010s
Extreme Arid Climate	T $\nearrow$	2010s
North American Blizzard	T $\searrow$	2011

It is evident that catastrophic climates are common in Wichita. We observe that various serious disasters which last longer than one year or with relatively large impacts tend to strike two or three times over the 100-year period, together with a high degree of randomness in time and unclear cyclical pattern, but descend more frequently as time goes by. Meanwhile, small-scale hot waves or other anomalous climates sustaining fewer than 3 months or causing a relatively small impact are relatively evenly distributed. Thus, we randomly insert 2 large droughts or floods and 2 severe heat or cold waves, together with 10 small droughts or floods and 10 hot waves or cold waves into our prediction for the next century to evaluate the communities' response constructed by our model.

As illustrated in Fig.17, orange dots represent the lasting time and severity of abnormal temperatures, as well as measurement of abnormal precipitation (in blue). Our modeled communities were almost stable under the impacts of continuous, large, and small weather anomalies. That is because the intricate symbiotic and synergistic relationships

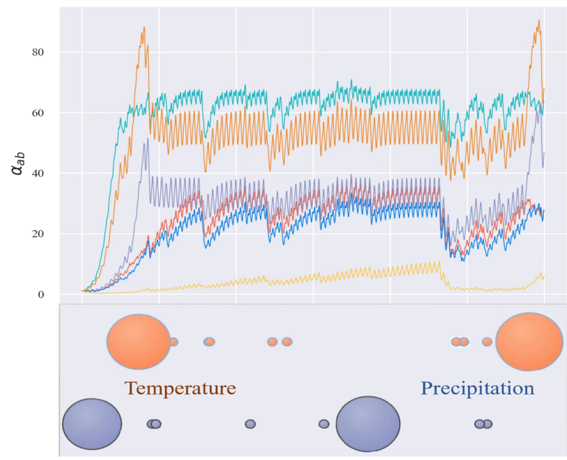


Figure 17: Community under more frequent and more various disasters

within a well-membered community can effectively protect vulnerable species from complete extinction, thus avoiding a positive feedback avalanche in the development level of the whole community caused by the disappearance of individual populations. For example, S_4, which once tended to become extinct under the great crisis shown, is significantly more affected by other organisms when its own biomass is at a low level, according to the principles of our model construction (indicated in 5.2). So with extensive symbiotic relationships, S_4 survived the catastrophe.

6.2 Influences of Other Factors on Results

6.2.1 Pollution

Pollution has many aspects, the majority of which are listed in the following part.

$\hookrightarrow$ Soil and water pollution

• The heavy metals elevate the intracellular levels of reactive oxygen species (ROS), which causes protein denaturation, including enzymes for photosynthesis, which may ultimately kill the entire plant. As an example, Cu can destroy the procedure of photosynthesis or even harm the whole plant at the concentration of 5 mg/kg. [8]

↪ Pesticide pollution

• The pesticide residues will have some deleterious effects. For instance, insecticides will block the stomata of plants, stopping the absorption of CO_2, then alterations in the photosynthetic process. They may also poison insects pollinating them, so their reproduction will be blocked, descending their growth rate. [8]

↪ Genetic pollution

• Contaminated altered genes from genetically engineered organisms may disperse to nature, causing irreversible pollution. For instance, if the most adaptive species S_6 receive the gene from humid areas, its drought resistance will be crippled, causing the reduction of species in the whole community. Homogenization can also take place, harming genetic diversity. [9]

↪ Species pollution (Biological invasion)

• If invasive plants occupy the community, the biodiversity will certainly decrease, reducing the resistance stability, which harms the long-term viability. When other creatures invade, they may eat up excessive plants in the community or herbivorous creatures. The former causes direct harm to the number of species, the latter can destroy the food chain and ultimately make plants die from starvation. [10]

Among these pollutions above, we choose soil and water pollution and genetic pollution as two symbols. We impose the two pollutions respectively on the community after all α, the biomass of species, reach stable. The results are as predicted:

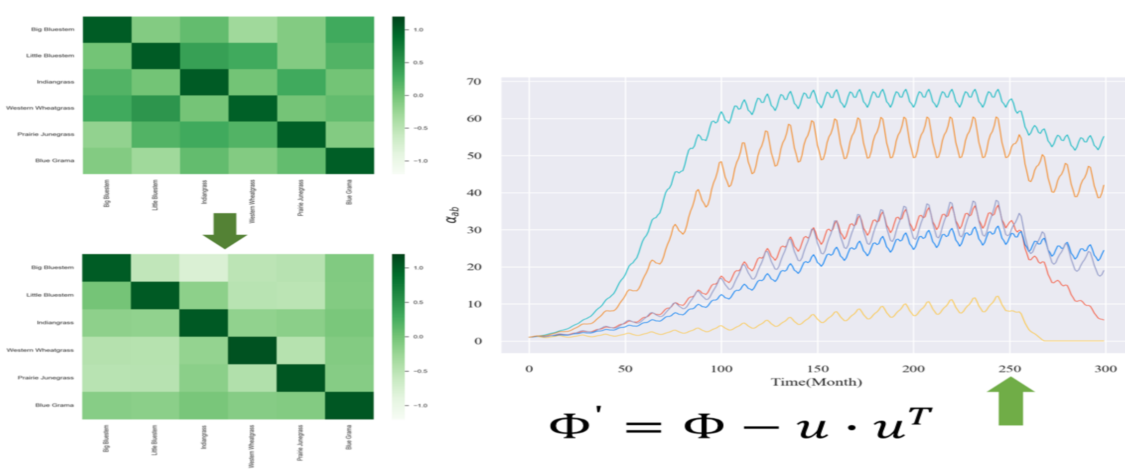


Figure 18: Community under the soil and water pollution

At the right of Fig. 18, the community is poisoned by the heavy metal Cu at the moment targeted, which blocks photosynthesis. It can be seen that the biomass of all species drops immediately and almost reaches 0 after 400 months, announcing the death of this community.

In addition to the harm of Cu itself, the protection effect of biodiversity is also weakened, as is shown in the left part. The two heat maps represent the variation of the correlation matrix Φ which measures the extent of interactions between species. It is notable that the value of all elements declines significantly. As a result, the whole community dies after 400 months of heavy metal pollution.

When it comes to genetic pollution, two species die out after about 70 months, with others returning stable. We can learn from the heat map that the maximum in Φ does not change at all, while the value of most of the elements declines. It can also be proved that genetic pollution damages the genetic diversity of biodiversity, as the colors in Fig. 19 become more monotonous.

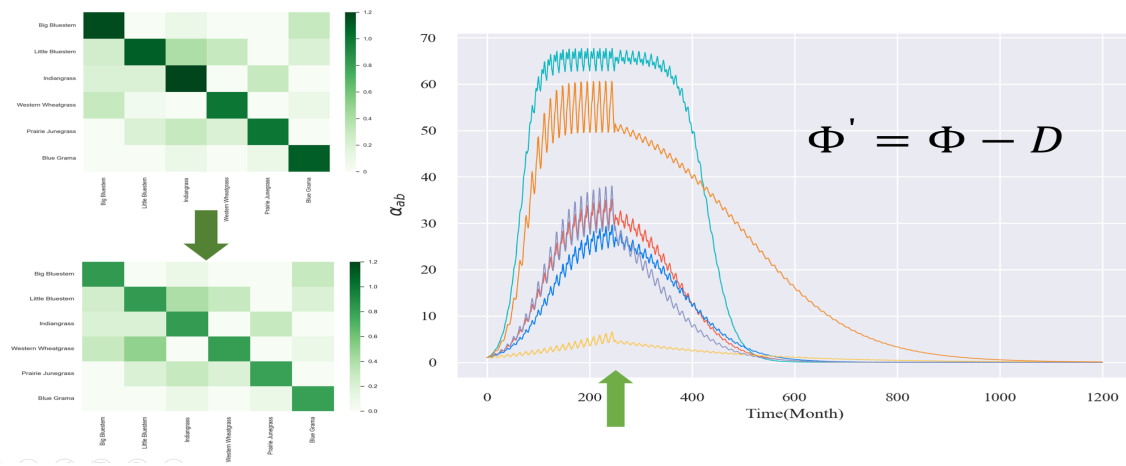


Figure 19: Community under genetic pollution

To conclude, pollutions have serious negative impacts on the plant itself, on the species, on the whole community, and on biodiversity.

6.2.2 Habitat Reduction

As our global assumptions show, we define the habitat of our specimens, which are six herbaceous plants, as a weather-evenly-conditioned circle area with a radius of R. On condition of overgrazing or reclamation, the radius will shrink to R', which can be found below:

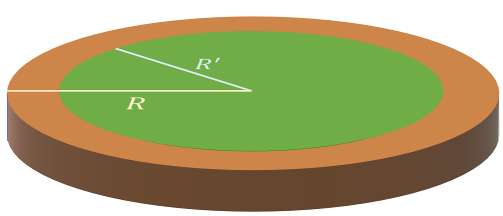


Figure 20: Shrinked habitat

According to the data from the work of M.He's team [2], the resource for plants is in proportion to the reciprocal of its area of habitat. Consequently, as the radius of the habitat shrinks, the elements of our correlation matrix Φ should scale with a ratio γ:

$$\gamma = \frac{S}{S'} = \frac{\pi R^2}{\pi R'^2} = \left(\frac{R}{R'}\right)^2 \tag{1}$$

Algorithm 1 algorithm of scaling the Φ matrix

Input: γ, Φ^{old}
Output: Φ^{new}

for each $i \in [1, n]$ **do**
  for each $j \in [1, n]$ **do**
    if $\Phi_{i,j}^{old} > 0$ **then**
      return $\Phi_{i,j}^{old} \leftarrow \frac{1}{\gamma}\, \Phi_{i,j}^{old}$
    else if $\Phi_{i,j}^{old} < 0$ **then**
      return $\Phi_{i,j}^{old} \leftarrow \gamma \cdot \Phi_{i,j}^{old}$
    else
      return $\Phi_{i,j}^{old} \leftarrow \Phi_{i,j}^{old}$
    end if
  end for
end for

And our result can be seen from the graph below.

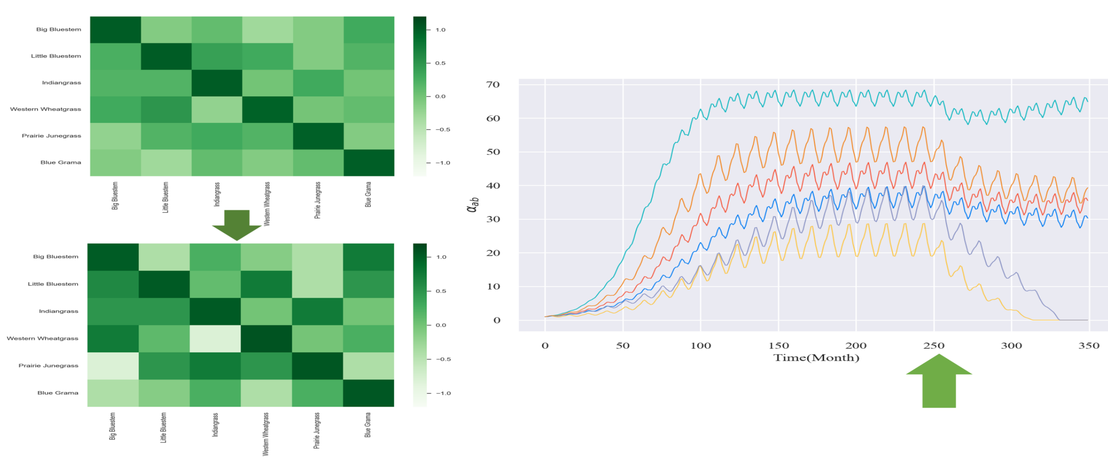


Figure 21: Community under habitat reduction

6.3 What Should We Do? What Is the Impact?

↪ As mentioned in 7.2, our model shows a number of biodiversity-harmful factors including pollution and habitat reduction mainly stem from anthropogenic activities, which ultimately destroy the whole community. It is thus necessary that we restrain them.

• **Control the use of fertilizers and pesticides.** This measure will ease the heavy metal pollution in soil and water, while preventing soil from compaction and relieving bioenrichment in animals and humans in a larger environment.

• **Be more careful for the sake of transgenic technology.** Conducting restrictive measures when raising transgenic creatures will control genetic pollution, thus protecting genetic diversity and preventing further damage to the ecosystem in a larger environment, such as the emergence of "super weed".

• **Fully investigate before bringing in foreign species.** Ensuring that the enemies of alien creatures exist attaches great security to the stability of ecosystems, as well as guard rare species and biodiversity in a larger environment.

$\hookrightarrow$ To approach the core of strengthening the long-term viability of plant communities and biodiversity, we should reduce the occurrence of irregular weather.

• **Extend forests.** As lungs of the earth, forests are indispensable in the water cycle and the regulation of climate. They also contribute to wind-break and sand-fixation and enrich biodiversity themselves in a larger environment.

• **Reduce carbon emission.** Controlling carbon emissions can relieve the greenhouse effect, which reduces the occurrence of hot waves and slows down the rise of sea levels.

7 Model Validation and Sensitivity Analysis

On a global scale where the different species' optimal temperatures and precipitations vary from one to another. If we want to extend our model to a more complex ecosystem, identify the dependency of our model on these initial conditions, which is to say, the changes of the α caused by the perturbation of species' optimal temperatures and precipitations. So here we take a range:

$$T_0 \in [54.5775, 60.3487] \quad , \ P_0 \in [1.6194, 1.7890]$$

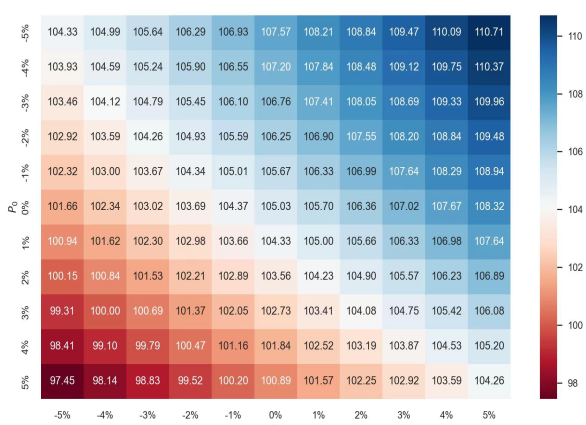


Figure 22: The sensitivity of our model under perturbation

From this graph above, we are able to observe that with other parameters and initial values fixed, when optimal temperature increases around our metric, the αre will monotonously increase, and the situation is approximately the same for optimal precipitation. Therefore, to some extent, our model is robust when considering the change in these two parameters, and thus indicates the possibility to apply our model to other ecosystems.

8 Model Evaluation and Further Discussion

8.1 Strengths

• **Calibration and normalization:** We utilize the Gaussian function to normalize the distribution of the exterior temperature/precipitation in model I and we construct a mapping function satisfying constraints and boundary conditions to quantify the environmental impact. Our calibration and normalization work unifies metrics in high dimensional realities and reduces the risk of round-off error in computation.
• **Validity:** Our model is based on the predecessor work of G. F. Gause [4] and all of our data are acquired from authoritative constitutions [3,11]. In model II, before adopting the ARIMA method, we do the ACF and PACF tests to make sure that it is suitable for our condition. What's more, after using the least square (ELS, OLS) method, we evaluate its error and find it relatively much lower than the acceptable criteria.
• **Novelty:** We have illustrated the correlation of different species in a community with the help of linear algebra, pairplot(each correlation element varies with time) and heatmaofnapshot of a state), which gives a multi-level vivid interpretation of our model.
• **Generalizability:** In sensitivity analysis, we found our model is quite robust to the optimal temperature and precipitation characteristics of different species, so aside from our case study, in a more complex ecosystem, our model will make sense.

8.2 Weaknesses

• **Model limitations:** Our six-specimen-community model, since is built on data from Kansas City, one region of the Great Plains. Thus it has some limitations and cannot be extended to different climates around the world. Additionally, this model is a comprehensive model for a community, which cannot reflect individualized activities.

8.3 Further Discussion

When we face ecosystems with far more species and much higher layer complexity, the calculation of our matrix became computationally unsolvable, and thus we need a more powerful approach to our hidden dynamic system. DMD (Dynamic Mode Decomposition) [12] method is a data-driven decomposition method for time series (recently only confines to the application of diagnostics and fluid mechanics), The algorithm is used for reducing dimensionality by detecting oscillating modes and determining their spatial structure and temporal behavior. Our team will further our study in this field and devote ourselves to mathematics modeling in ecology.

References

[1] 1-pick-drought-tolerant-plants. 2012, February7. [Online]. Avail-able: https://powerofplants.com/2012/02/7-steps-to-a-desert-wise-landscape/1-pick-drought-tolerant-plants/

[2] M. He, Y. Pan, G. Zhou, K. E. Barry, Y. Fu, and X. Zhou, "Grazing and global change factors differentially affect biodiversity-ecosystem functioning relationships in grassland ecosystems," *Global Change Biology*, vol. 28, no. 18, pp. 5492–5504, 2022.

[3] FOURTH NATIONAL CLIMATE ASSESSMENT. [Online]. Available: https://nca2018.globalchange.gov/downloads/

[4] G. F. Gause, "Experimental populations of microscopic organisms," *Ecology(Ecology)*, vol. 18, p. 173–179, 1937.

[5] Definition of 'biomass'. [Online]. Available: https://www.collinsdictionary.com/dictionary/english/biomass

[6] D. Lobell, K. Nicholas, and C. Field, "Historical effects of temperature and precipitation on california crop yields," *Climatic Change (Clim. Change)*, vol. 81, pp. 187–203, 2007.

[7] Definition of 'allelopathy'. [Online]. Available: https://www.collinsdictionary.com/dictionary/english/allelopathy

[8] A. Alengebawy, S. T. Abdelkhalek, S. R. Qureshi, and M.-Q. Wang, "Heavy metals and pesticides toxicity in agricultural soil and plants: Ecological risks and human health implications," *Toxics(Toxics)*, vol. 9, no. 3, 2021.

[9] P. Patel, T. Desai, Y. Virdiya, and K. Kugashiya, "Genetic pollution in plants: A review," 2022.

[10] D. M. Lodge, "Biological invasions: Lessons for ecology," *Trends in Ecology Evolution(Trends Ecol. Evol.)*, vol. 8, no. 4, pp. 133–137, 1993.

[11] C. R. R. I. Latham, J. and M. Bloise, "Global land cover share (glc- share) database beta- release version 1.0– 2014," p. 29, 2014.

[12] J. N. Kutz, S. L. Brunton, B. W. Brunton, and J. L. Proctor, *Dynamic mode decomposition: data-driven modeling of complex systems*. SIAM, 2016.

[1–12]