IMPERIAL

Imperial College London

Utilizing Hill Numbers to Analyze Species Evenness Time Sequences Across Multiple Dimensions:

A Study Based on the Integration of Living Planet Index and Global Abundance Data for 9700 Birds

Author: Fanshu Xu

Supervisor: Prof. James Rosindell

A thesis submitted in partial fulfilment of the requirements for the degree of Master of Research at Imperial College London

Submitted for the MRes in Computational Methods in Ecology & Evolution

August 2024

Contents

Declaration

All data used in this research is publicly available, with the exception of the LPI completeness data, which has not yet been formally published. The datasets mainly originate from the Living Planet Index (LPI). I take full responsibility for the data processing, analysis, and visualization. My supervisor, James Rosindell, offered detailed and patient guidance throughout the project.

Abstract

The conservation of biodiversity has garnered significant attention in recent years. Quantifying species diversity provides a robust theoretical foundation for formulating biodiversity conservation policies. Hill number and living planet index(LPI) are both robust indicators for community's development. In this research, I aimed to comprehensively elucidate species diversity trends of birds in global scale by integrating Hill numbers with a weighted LPI by global abundance. I estimated global abundance development by combining a global abundance dataset of 9,700 birds with the dataset about species growth generated during the Living Planet Index calculation process. Furthermore, I conducted analysis on abundance distribution based on combined dataset and calculation of evenness sensitivity based on Hill number. This multifaceted approach offers a more nuanced understanding of biodiversity dynamics especially in terms of species evenness, potentially informing more complete global biodiversity monitoring where data are available.

Keywords: Biodiversity, Hill Numbers, Living Planet Index, Species Evenness, Global Abundance

1 Introduction

The biological diversity of ecosystems has always attracted the attention of environmentalists and ecologists (Wilson 1995). In recent decades, global biodiversity has been decreasing and extinction rates of species are much higher than background rates (Chaytor et al. 2002). Key biodiversity indicators such as species' population trends, extinction risk, habitat extent and habitat condition show a downward trend (Butchart et al. 2010). The Convention on Biological Diversity (CBD) illustrates targets for a global effort on biodiversity conservation, sustainable utilization, and equitable benefit-sharing (Rosendal 2000). The objective of CBD in 2010 is to preserve ecosystem services, uphold the well-being of the earth, and guarantee that crucial advantages are accessible to all individuals and the environment(CBD 2010). Most countries, however, failed to meet the targets of CBD's 'Strategic Plan for Biodiversity 2011–2020'

(Buchanan et al. 2020). Then, CBD's vision for the future was articulated, emphasizing that biodiversity should not only be valued but also conserved, restored, and utilized in a sustainable manner.(Joly 2022). Biodiversity is the sum total of all biotic variation from the level of genes to ecosystems (Purvis & Hector 2000). Biodiversity is a broad and complex concept that is difficult to quantify directly, so researchers often use appropriate surrogate indicators when analysing biodiversity(Sarkar 2002). Species diversity is an indicator that is relatively easy to qualified and its main factors are species evenness and species richness. The richness is a measure of the number of different species present in a particular area or ecosystem while species evenness originally refers to the uniformity of the abundance distribution among different populations within a community. Even when species richness remains unchanged or changes only a little, evenness can fluctuate where the abundances of dominant species fluctuate significantly over time. Evenness has been shown to exhibit a significant linear relationship with total biomass when richness level remains nearly constant over time(Wilsey & Potvin 2000). Richness and evenness are both regarded as important indicators for diversity at first but species richness has been more frequently used among diversity indicators because of its high accessibility to the public and to policy makers and researchers often restrict consideration of diversity to only species richness (Andelman & Willig 2003). A research about principal components analysis(PCA) for different biodiversity indexes showed that different indices including richness, evenness performed distinctively on two main principal components of diversity in different sampling sites after dimension reduction, which means that these indicators cover different aspects and actually, species richness could not fully represent the biodiversity(Wilsey et al. 2005). So it has been suggested that species evenness is a main factor of species diversity as well and the indices for species richness should be combined with it to represent the effective diversity(Tuomisto 2010).

To monitor progress towards the CBD's 2050 goals, it is important to put forward appropriate quantitative measures of global biodiversity(Chandra & Idrisova 2011). Hill number is an excellent index for measuring effective species diversity that is widely used in ecology. It can take into account both the species richness and evenness of certain community and can adjust the sensitivity of species diversity to the relative abundance of species by adjusting the q value according to the characteristics of the species composition in the communityChao et al. (2014). Shannon Index and Simpson Index are both widely used types of diversity measurement and Hill number is exponential result of Shannon entropy when q value equals 1 while it is the inversion of Simpson Index when q equals 2 (Keylock 2005). It could generally represent diversity rather than focusing on a single factor concerning diversity. Regardless of how the p-value is adjusted, the Hill number remains influenced by species richness during the calculation process and can only capture changes in species evenness(Mouillot & Wilson 2002). The calculation of Pielou evenness is more independent with species richness(Gosselin 2006) and it is based on the Shannon index as well. Hill number is a practical and robust measurement of species diversity, but constructing diversity time sequences via Hill numbers requires relative abundance data for most species in the group over time, which is often difficult to obtain. As a result, there is a lack of data to enable Hill number analyses through time and study change at the global scale.

The Living Planet Index (LPI) is another widely utilized resource in biodiversity studies. The Living Planet Index (LPI) is an index with associated metrics that tracks population trends across numerous species in diverse ecosystems. Its methodology involves calculating the growth rates of each population's abundance over time, referred to as lambda values. These lambda values represent the growth multiple of each population based on the Living Planet Dataset (LPD). So LPI is obtained based on solid data foundation concerning abundance time sequences of large number of populations in global scale. First, lambda values are calculated for each population, and then averaged across all populations of the same species within each realm after standardization. The time sequences of species lambda values in different realms are then used as exponents in a cumulative multiplication calculation, starting from 1 in the base year, to derive the relative index sequences for each realm and system. LPI is used as an index to illustrate the general trend in changes in the community's total biomass, which in turn indicates potential changes in

species diversity (Loh et al. 2005). Consequently, LPI has become widely accepted because it is simple to calculate and covers abundance time sequences of global populations of many species.
Relative abundances of species are a critical factor in evenness (Tuomisto 2010). In addition, total biomass of a community has been shown in some cases to have a significant positive correlation with its diversity(Bock et al. 2007) and it is determined by species richness and their abundances. As an indicator for species diversity, species richness in certain is simple to calculate but it disregards the effect of abundance(MacArthur 1965). As previously mentioned, species abundances are closely related to the primary factors influencing species diversity concerning species evenness and community size. The Living Planet Index (LPI) indirectly links to species diversity through the overall trend in community's biomass. Thus, it is beneficial for the LPI to incorporate relative abundance data for each species to more accurately reflect their contributions because the biomass has direct linear relationship with species abundance (Hallmann et al. 2021). This could be achieved by calculating weighted average species lambda time sequences based on relative abundance. A previous research also showed that employing appropriate indicators for weighting when calculating the LPI can somewhat address the shifting pattern of biodiversity(Buschke et al. 2021). If this weighting process is applied, there is a hypotheses remain to be tested whether species abundance has significant positive relationship with its growth potentiality. Moreover, the LPD uses 2476 units for species's abundances corresponding to various methods used by its sources for collecting data and converting between different units can be extremely challenging. As a result, there is no unified unit for abundance and no effective weighting by relative abundance can be done in the process of calculating the average of species lambda of in each realm and the species lambda sequences in each realm in LPI calculation are actually not comparable. This presents an opportunity to enhance the LPI by incorporating the effects of relative abundance. Addressing this aspect could improve the LPI's ability to more effectively reflect potential changes in effective biodiversity.
In this research, I referred to a dataset of 9,700 birds' global abundance which is based on e-bird dataset from 2010 to 2019(Callaghan et al. 2021)

and combined it with species lambda dataset to create a dataset of estimated global abundance time sequences for birds so that I could solve the problem that there is no consistent unit of abundance in LPD. This combined dataset serves as a foundation for calculating both a weighted Living Planet Index (LPI) based on relative abundance and Hill numbers. Although the species lambda dataset does not indicate any extinctions and bird richness levels have remained relatively stable across time, multi-dimensional analysis of species evenness can still be conducted derived from Hill number calculations.

2 Methods

All of the data analyses are accomplished using R version 4.3.1(R Core Team 2023). Generally, I conducted data cleaning for species lambda dataset in order to raise its effectiveness. This birds' abundances dataset is in global scale, so it should be analysed together with dataset of species lambda time sequences so that I could get species global abundance times sequences. After that, I calculated weighted LPI by abundance and Hill number. Moreover, I conducted relative evenness calculation and evenness sensitivity analysis based on the result of Hill number calculation. The overview of research flow could be found in (Figure 1).

2.1 Global Species Lambda Calculation

The growth rate sequence of each population is calculated firstly to prepare getting species lambda sequences.

$$d_{ijt} = \log_{10}\left(\frac{N_{ijt}}{N_{ij(t-1)}}\right) \tag{1}$$

The d_{ijt} value stands for lambda value of the j-th population of the i-th species, which means its growth rate in year t and the N_{ijt} value stands for the abundance value in this year for j-th population in LPD.

$$\bar{d}_{it} = \frac{1}{m_t}\sum_{j=1}^{m_t} d_{ijt} \tag{2}$$

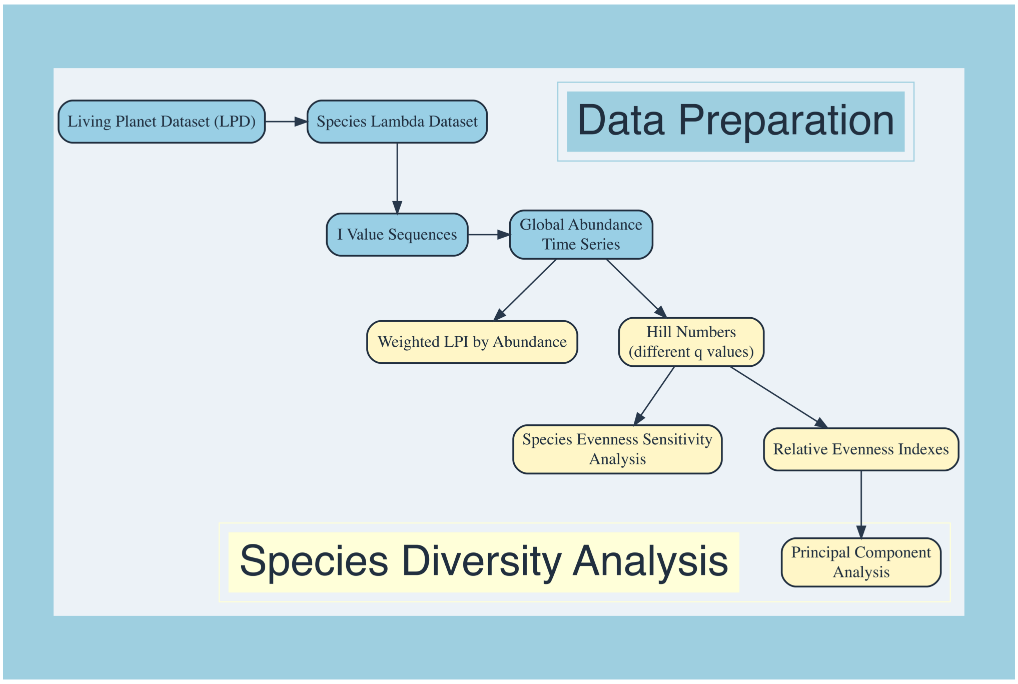

Figure 1: Flowchart of the project including two periods concerning data preparation and species diversity analysis.

The d_{it} values stands for lambda values for certain species in year t, which is the arithmetic mean of lambda values of all populations for this species in this year and the m_t stands for the number of populations of the i-th species of which the lambda value in year t is not null.

In the calculation of growth rates, this chain method is frequently employed for cumulative calculations (Loh et al. 2005).

2.2 Data Cleaning

These selection standards including the proportion of null values, the representative of all populations in global scale according to the percentage values in LPI completeness data and different species' global abundance sizes.

Aiming at filtering species lambda sequences to find those with less null values and determining the starting and ending time points for all sequences, I computed the proportion of null values separately for each species sequence and time column in the dataset of species lambda. After fitting a normal distribution, I retained the species (rows) and year (columns) in species lambda dataset with proportions of null values below the mean.

Additionally, I made another selection of species lambda dataset to identify species encompassing populations that could comprise more than 25% of the their global distribution in accordance with proportion values in LPI completeness data. This adjustment aims to mitigate biases and enhance the representation of the original dataset.

Moreover, I created five distinct abundance-based species groups: those with global abundances above the top 25%, above the top 50%, below the bottom 50%, below the bottom 25%, and within the central 50%. These groups correspond to the dominant species group, less dominant species group, less rare species group, rare species group, and middle species group, respectively. This selection was intended to provide a data foundation for calculating LPIs, Hill numbers, and other indices for both common and rare species.

2.3 Data Matching

Firstly, I have made a correction for birds' scientific names in both of them using the rotl package(Michonneau et al. 2016) in order to match LPD and birds dataset better based on the scientific names of birds in them.

$$I_{it} = I_{i(t-1)} \times 10^{d_{it}} \tag{3}$$

The I_{it} value stands for the I value of the i-th species in year t. The I value basically starts from 1, determined by I value and species lambda value in previous years. So I value sequence follows cumulative calculation. I computed the I value sequence for each species and subsequently determined the geometric mean of values across all time points. The I value sequence is cumulative calculation of abundance trends and the ratio between averaged I values and global abundance shows the disparity in magnitude between them. I applied a multiplication factor to the annual values within the respective I sequence, by utilizing the aforementioned ratio to estimate a global abundance time sequences for each bird that are necessary for calculating Hill numbers (Jost 2006). In addition, all I values are positive, making it possible for conducting this data scaling for all species.

$$\text{Averaged I-i Value} = \exp\left(\frac{1}{n}\sum_{n=1}^{n}\log(I_{i_t})\right) \tag{4}$$

The n values stand for the number of time points.

$$\text{Abundance}_{it} = I_{i_t} \times \left(\frac{\text{Global Abundance}}{\text{Averaged I-i Value}}\right) \tag{5}$$

2.4 Abundance Distribution Analysis

First, I applied a base-2 logarithmic transformation to the obtained global abundance sequences, resulting in a time series of abundance octave data(Hubbell 2001). Moreover, I calculate the standard deviation of different abundance-based species groups' abundance octave distribution in each year to show the change of evenness level in one aspect.

2.5 Original LPI Calculation

The average d_t and I_t time series for all of these populations were computed, following the standard procedure for calculating LPI values.

$$\bar{d_t} = \frac{1}{n_t}\sum_{i=1}^{n_t} d_{it} \tag{6}$$

In the computation of the original LPI, an arithmetic average was employed to determine the mean d_t value. The d_{it} is the growth rate of the i-th species in year t. The n_t stands for the number of not null average species lambda values in year t.

$$I_t = I_{t-1} \times 10^{\bar{d_t}} \tag{7}$$

A cumulative multiplication operation is then applied. The outcomes of calculation are mainly determined by the average species lambda in each year. The I_t values are actually relative values initiating with a standard value of 1. In the step of data matching, the I value sequence would show the trend in the change of abundance of certain population while in the calculation of original LPI, this cumulative calculation is applied to present overall trends in total biomass of certain groups of species.

2.6 Weighted LPI Calculation

In the weighted LPI calculation, the weighting process involves calculating the weighted average of species' lambda values. I proposed a hypothesis regarding the biological significance of this weighting process and tested it by fitting a generalized linear model (GLM).

The metabolic rate of certain species is strongly correlated with both body size and ambient temperature (Brown et al. 2004). This research of brown also demonstrates a significant linear relationship between metabolic rate and individual biomass production. Moreover, the maximum growth rate of a species is typically scaled by its metabolism, as reproduction is driven by metabolic processes. Finally, I used the R function 'glm' to establish a relationship between maximum species lambda value with global abundance data and other factors in my research, guided by insights from previous research on metabolic theory. To incorporate metabolic information into species lambda sequences, I used the maximum species lambda value as a proxy for metabolic rates. I also used the arithmetic mean of latitude data from global populations in the LPD to represent temperature factors. Additionally, I employed global abundance development to represent biomass production and referenced to species body mass data(Morales-Castilla et al. 2012) to link the obtained variances with corresponding variances in metabolic theory.

The weighting process in averaged species lambda calculation followed another formula. The P_{i_t} value represents the relative abundance data of the i-th species in year t.

$$\bar{d}_t = \sum_{i=1}^{n_t} d_{it} \times P_{i_t} \tag{8}$$

2.7 Effective Diversity Calculation

2.7.1 Hill number(q=1) Calculation

$$^{1}D_t = \exp\left(-\sum_{i=1}^{S} P_{i_t} \ln(P_{i_t})\right)$$

The $^{1}D_t$ values represent the exponential form of Shannon Index values in year t while the P_{i_t} values denote the relative abundance of the i-th species in year t.(Bromiley et al. 2004)

2.7.2 Hill number(q=2) Calculation

In my research, species richness almost remain unchanged during given period of time in each abundance-based species group, so evenness has large influence on the change of diversity. Therefore, I calculated Hill numbers with q values of 1 and 2 separately, aiming to provide a foundation for analyzing species evenness while also quantifying species diversity across different abundance-based species groups. $^{2}D_t$ value represents the inversion of Simpson Index value in year t.

$$^{2}D_t = \sum_{i=1}^{s} P_{i_t}^{2} \tag{9}$$

2.7.3 Evenness Indexes Calculation

Pielou evenness is an robust evenness index that its calculation is relatively independent from species richness(Jost 2010). This method aims to assess the extent to which the distribution of species abundances deviates from perfect evenness, thereby serving as the relative evenness index of certain community(Studeny et al. 2011). In the calculation of Pielou evenness, H_t values refer to Shannon Index value in year t and J_t values represent the Pielou evenness in year t while S_t values stand for species richness level in this year.

$$H_t = -\sum_{i=1}^{S} P_{i_t} \ln(P_{i_t})$$

$$J_t = \frac{H_t}{log(S_t)} \tag{10}$$

There is another index of evenness based on Hill number and have similar mathematical principles to Pielou evenness. A previous research proposed a new method to measure relative evenness based on the likelihood ratio G, Pearson's X2 and Hill number(Cressie & Read 1984, Read & Cressie 2012).

$$I_{p_t}(\lambda) = \frac{1}{\lambda(\lambda+1)}\left[\left(\frac{N(0)_t}{N(\lambda+1)_t}\right)^{\lambda} - 1\right] \tag{11}$$

Here, $I_{p,t}(\lambda)$ serves as an index to measure the degree of deviation, where higher values indicate lower levels of relative evenness. The $I_{p,t}$ value represents the relative evenness level in year t, while λ refers to the weight assigned to dominant species. $N(0)_t$ and $N(\lambda+1)_t$ represent the Hill numbers in year t when q equals 0 and $\lambda+1$, respectively.

2.8 Evenness Sensitivity

I calculated growth rate at each time point of each Hill number sequence and I measured the degree of changes in grow rates when the q value increased from 1 to 2 by calculating the absolute value of ratio between the changing rates of the Hill number at q = 1 and that of the Hill number at q = 2 at each time point. This indicator shows how the ability of species evenness to influence species diversity would change when the weight of richness decreases and that of evenness grows in each abundance-based species group, which means the evenness sensitivity to species evenness.

$$\text{Evenness Sensitivity}(t) = \left|\frac{\frac{d}{dt}{}^{2}D_t}{\frac{d}{dt}{}^{1}D_t}\right| \tag{12}$$

2.9 Analysis on Different Dimensions of Evenness

I calculated arithmetic mean of evenness sensitivity, the arithmetic mean of standard deviation in different abundance-based species groups' abundance distribution and arithmetic mean of these groups $I_p(\lambda)$ in all year range. When calculating both $I_p(\lambda)$ and Pielou evenness time sequences, I largely mitigated the influence of species richness on evenness. However, the potential ecological interaction between evenness and other biodiversity factors remains uncertain. To determine if ecological connections exist between species evenness and other factors of species diversity, and to identify the specific dimension of evenness where these connections may reside, I conducted Generalized Linear Model(GLM) fitting on different abundance-based species groups' $I_p(\lambda)$ with their SD of abundance distribution and evenness sensitivity over time.

3 Results

3.1 Overview of Changes in Abundance Distribution

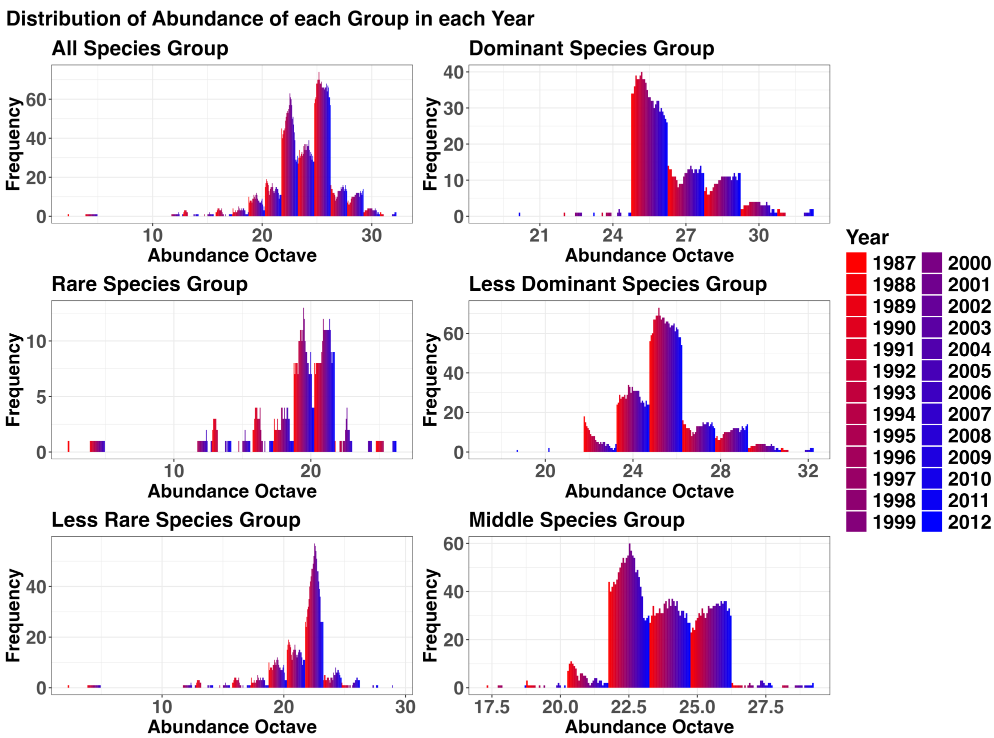

Figure 2: The frequency distribution of species across different abundance octaves, based on the logged abundances within various abundance-based species groups for each year. Different colors represent different years.

In six different abundance-based species groups, the abundance distribution of each group changes greatly over time(Figure 2) and each abundance-based species group mainly consists of proportionally more common species compared to previous studies (Loh et al. 2005, McRae et al. 2017). Even in the rare species group, the abundance octave with most species is over 20.

3.2 Hypothesis Test about Weighted LPI by Abundance

$$\begin{aligned}\text{Maximum Growth Rate} = &-0.168641\\ &+0.021672\times\log(\text{Body Mass})\\ &+0.018891\times\log(\text{Global Abundance})\\ &-0.002928\times\text{Latitude}\end{aligned} \tag{13}$$

	Estimate	Std. Error	t value	Pr(> |t|)
(Intercept)	−0.168641	0.091022	−1.853	0.064901
log(Body.Mass.gram.)	0.021672	0.005148	4.210	3.38×10^{-5} ***
log(global_abundance)	0.018891	0.005312	3.556	0.000437 ***
Latitude	−0.002928	0.000640	−4.575	6.99×10^{-6} ***

*Significance codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1*
Dispersion parameter for gaussian family taken to be 0.03030405
Null deviance: 10.3391 on 302 degrees of freedom
Residual deviance: 9.0609 on 299 degrees of freedom
AIC: −193.58
Number of Fisher Scoring iterations: 2

Table 1: GLM Model Fitting Results

The residual deviance of the model, with 299 degrees of freedom, is 9.0609, suggesting a good fit for the GLM. The coefficient's p-value for the term log(global_abundance) is 0.000437, indicating a significant positive correlation between the max species lambda value and global abundance(Figure 4).

3.3 LPIs & Hill Numbers in Different Abundance-based Species Groups

In order to compare the overall trends of Hill numbers with LPIs in each abundance-based species group and measure degrees of their changes, I

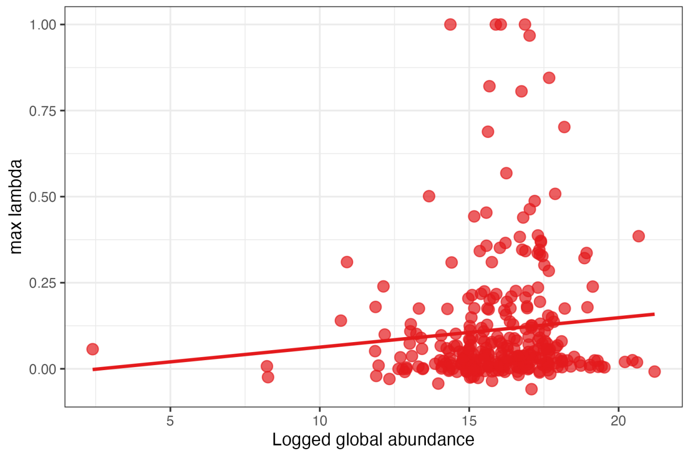

Figure 3: Hypothesis Test for Abundance's Influence on Species' Growth

scaled their index values so that the initial values are all set to 1, so values of each index are not comparable with each other except weighted LPI by abundance and original LPI. The divergence of these 2 LPIs is not very obvious in most species groups except for all species group and dominant species group. Furthermore, LPIs have quite different trends with other indexes in dominant species group. From the aspect of the development trends, Hill numbers with different q values generally exhibit a trend of initially rising and then falling in all abundance-based groups. Additionally, Hill number shows more dramatic changes when q equals 2 compared with Hill number(q=1).

3.4 Original and Weighted LPIs

The main difference between these two sets of lines lies in their degrees of fluctuation, with the weighted LPI lines exhibiting more pronounced volatility. Additionally, in the original LPI lines, growth rate values are positive in most years for all groups except rare species group. In contrast, the weighted LPI lines show a greater occurrence of negative growth rate values

3.5 Hill Number Sequences

When q equals 1, Hill number values of the middle species were the highest, while the minimum values appeared in index of the rare species group.

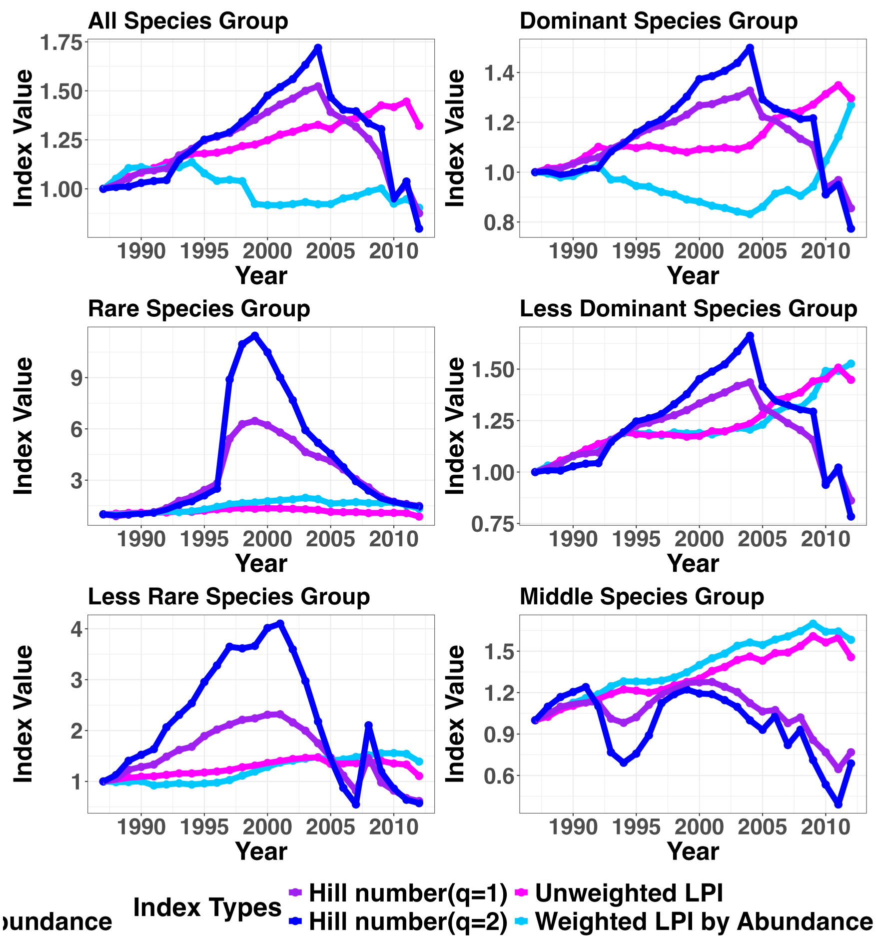

Figure 4: Line chart of different indexes sequences in different abundance-based species groups. Lines in different colors represent different types of indexes including weighted LPIs, original LPIs, Hill numbers(q=1) and Hill numbers(q=2).

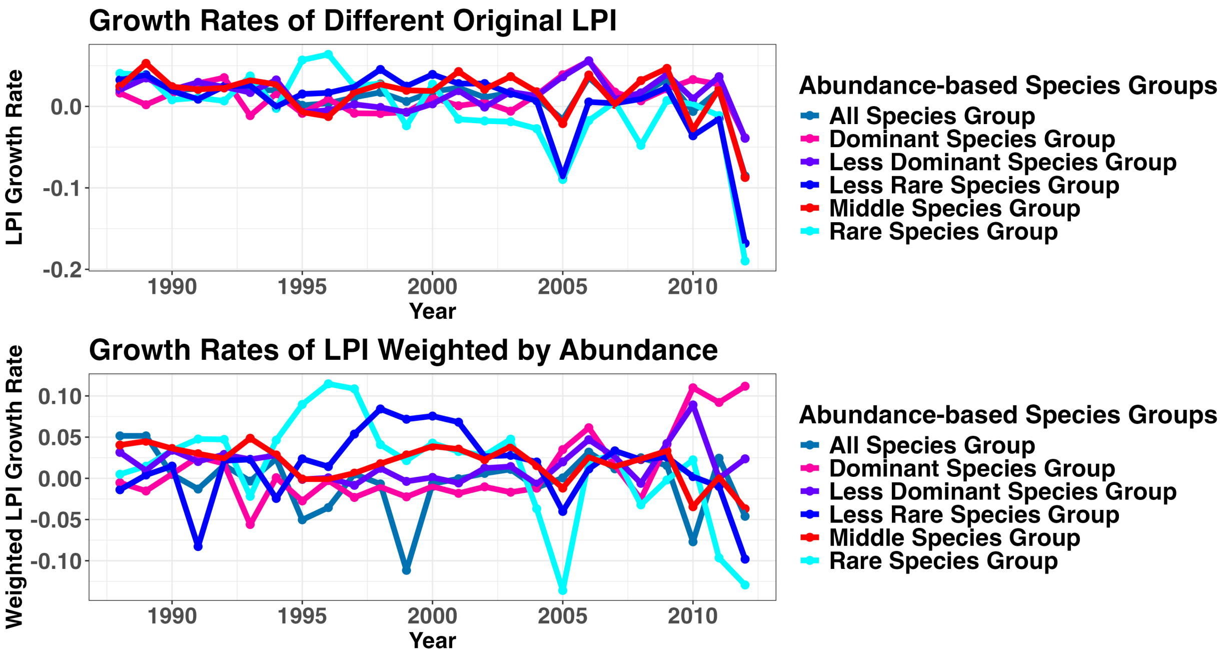

Figure 5: Lines of original LPIs' and weighted LPIs' growth rate time sequences.

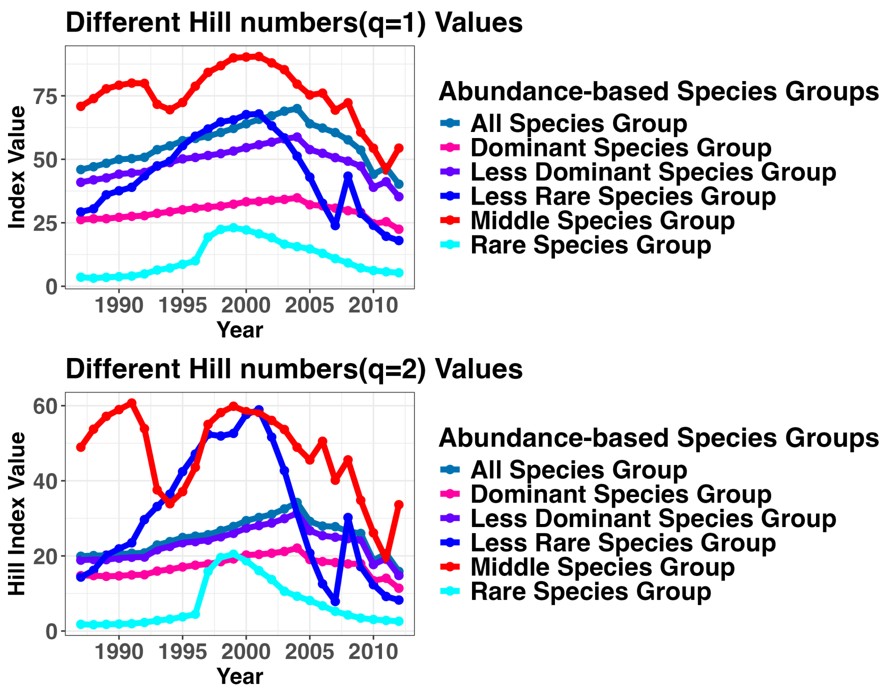

Figure 6: Top subplot: Lines of Hill number(q=1). x axis: Year. y axis: Index Value. Bottom subplot: Lines of Hill number(q=2).

The relative sizes of indexes values of six different abundance-based species groups change greatly after the q value of Hill number increased. When I increased the q value from 1 to 2, values of every sequence dropped by different degrees. Furthermore, the highest values and the second highest values exist in index of middle species group or less dominant species group. In terms of Hill numbers' development tendencies, sequences of middle species group or less dominant species group fluctuated more dramatically, whereas other sequences were relatively stable. In Figure 4, all Hill number sequences were scaled for comparison with LPI trends, meaning the displayed trends do not reflect their actual patterns. In Figure **??**, the Hill number trend was adjusted, resulting in much more stable sequences across all species groups, including dominant, rare, and less dominant species groups.

3.6 Relative Evenness Indexes

In terms of Pielou evenness, intense changes in index values only occur in the rare species group and the less rare species group. Values of Pielou evenness in the middle species group are the highest at most time points while the lowest values of Pielou evenness regularly exists in rare species group. The evenness rankings of the abundance-based species groups, as indicated by $I_p(\lambda)$, closely align with their rankings according to Pielou's Evenness Index. Moreover, general trends of Ip(λ) sequences with different λ values are similar to each other and fluctuations in Ip(λ) are much larger than those in Pielou Evenness in all groups Figure 7. Moreover, there is almost no change occurring in trends of Ip(λ) sequences when λ value changes.

3.7 Two Different Dimensions of Species Evenness

The abundance distribution serves as one dimension of relative evenness in this research while different abundance-based species groups sensitivity to evenness is regarded as another dimension of relative evenness. The standard deviation in abundance distribution across different abundance-based species groups does not follow the same trend with these groups' Ip(λ) sequences. The most even abundance distribution exists in domi-

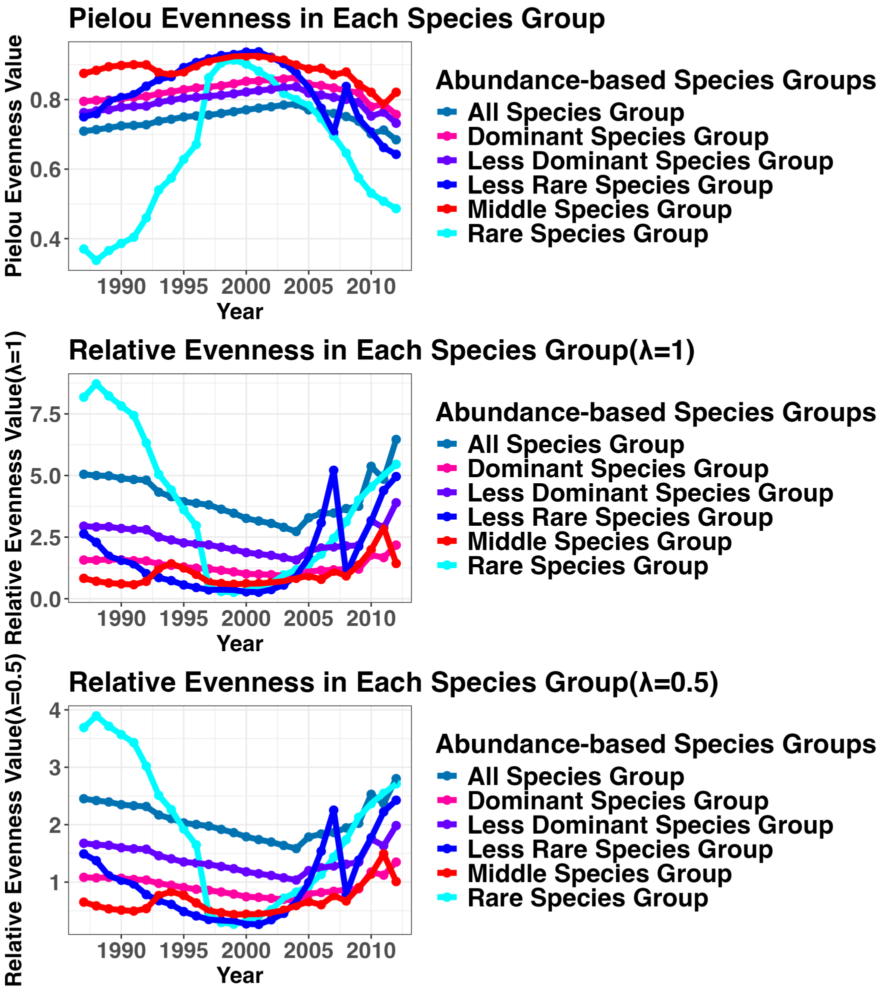

Figure 7: Top subplot: Lines of Pielou Evenness. Middle subplot: Lines of Ip(λ)(λ=1). Bottom subplot: Lines of Ip(λ)(λ=0.5).

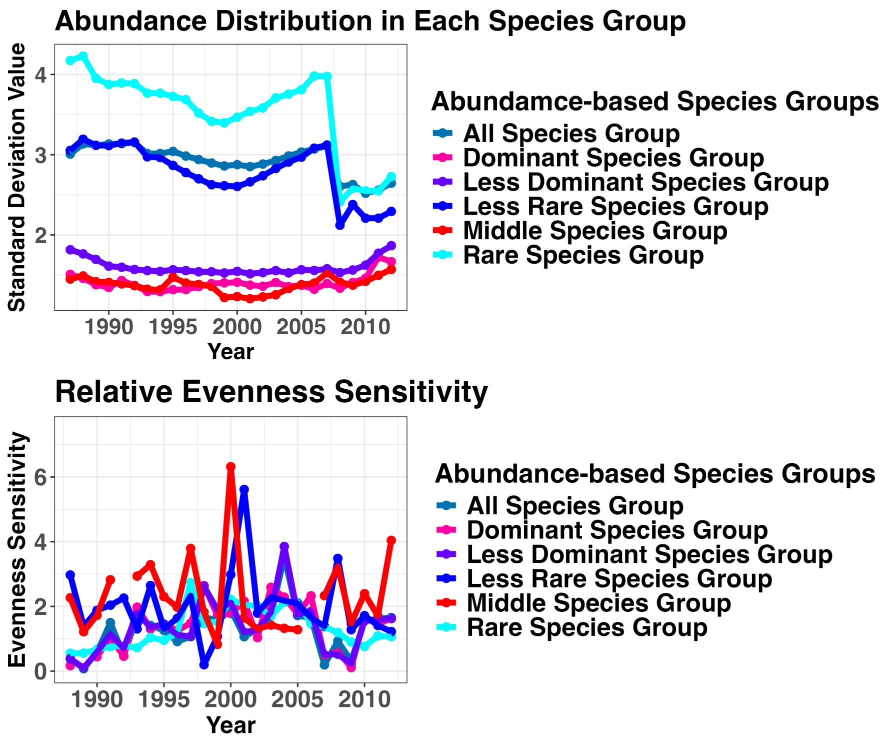

Figure 8: Top subplot: Lines of SD in Abundance Distribution in Each Abundance-based Group. Bottom subplot: Lines of Evenness Sensitivity.

nant species group, less dominant species group or middle species group. Concurrently, the highest sensitivity to evenness regularly exists in less rare species group at some time points(Figure 8). Results of the statistical analyses also indicate that average relative evenness, average SD of abundance distribution and average sensitivity to evenness over time do not follow the same trend within different abundance-based species groups. The Average standard deviation of less rare species group over time is large but its average sensitivity to evenness and average relative evenness level across time are relatively high.

Species Groups	**Average Ip(λ)(λ=1) Value**	**Average SD of Abundance Distribution**	**Evenness Sensitivity**
All Species	4.08	2.93	1.27
Top 25%	1.32	1.40	1.68
Bottom 25%	3.66	3.54	1.32
Top 50%	2.38	1.60	1.37
Bottom 50%	1.63	2.79	2.00
Middle 50%	0.99	1.38	5.72

Table 2: Two Different Dimensions of Relative Evenness Concerning Abundance Distribution and Sensitivity to Evenness in Different Abundance-based Species Groups

In the fitting outcome of GLM in different abundance-based species groups, two terms concern evenness sensitivity and abundance distribution. The fitting results indicate that the coefficient for abundance distribution is significant (with low p-values) in abundance-based species groups that have a small average standard deviation of abundance distribution. In terms of the coefficients of evenness sensitivity, all of the p-values do not reach significant level except the p-value for this coefficient in less dominant species group.

P-values for the coefficients of the various terms	All Species Group	Dominant Species Group	Less Dominant Species Group	Rare Species Group	Less Rare Species Group	Middle Species Group
Evenness Sensitivity	0.10	0.80	3.33×10^{-8}	0.13	0.17	0.53
Abundance Distribution	0.35	0.030	0.19	1.16×10^{-6}	0.45	0.022

Table 3: Outcome of GLM Fitting on Relative Evenness(Ip(λ)) with Sensitivity to Evenness and SD in Abundance Distribution across different abundance-based species groups

4 Discussion

4.1 Hypothesis Test about Weighted LPI by Abundance

The GLM results, which modeled maximum species lambda using species' global abundance, average latitude across populations, and body mass as predictors, show that the abundance factor has a significant impact on population growth. This suggests that weighting species lambda by global abundance may have biological relevance, as species abundance appears to positively influence growth potential, consistent with previous studies on metabolic theory(Brown et al. 2004).

4.2 Limitation and Future Work

During data preparation, I calculated the I value sequence for each species based on the selected species lambda dataset. All I value sequences begin at 1 and reflect the overall trends in species abundance over time. So the robust calculation of weighted LPI could also be achieved theoretically via calculating I value sequence of each species firstly and then getting weighted average of different species I values by global abundance. However, many species have null lambda values in the initial years, making it difficult to calculate I value sequences for numerous species. So future LPI research should focus on analyzing time sequences of population abundance in the Living Planet dataset using machine learning. This could involve referencing various variables in the LPD, such as populations' latitude data, to develop suitable Convolutional Neural Network (CNN) models for training original LPD and predicting null values in each population's abundance sequence.

4.3 Adjustment for LPI

The basic LPI calculation is based on one of the most extensive time series databases concerning vertebrate populations(Collen et al. 2009). In terms of my research, each abundance sequence I have obtained is considered as a combined population of global populations for each corresponding species because the species lambda sequences obtained via the rlpi package represent averaged growth rates series of all populations of each species, and the 9700 birds' abundance data covers a global scale as well(Tuomisto 2010). Therefore, the LPI calculation in my research is conducted at the level of species' global populations, rather than individual populations.

In previous studies, weighted averaging was also used in the calculation of original LPI, but not weighted via abundance. In the calculation of the average species lambda of each realm, the lambda time series of different species were firstly grouped by class, and then the data at each time point in each class group were averaged to obtain the average lambda time sequence of each class. After that, the lambda time sequences of all classes were weighted and averaged by the actual number of species in each class to obtain the weighted lambda time series by species numbers of each realm(McRae et al. 2017). The numbers of species in different classes serve as the contributions of species' growth in each class to the development of each realm. Given the scope of my research, the LPI study is only about aves because I matched species lambda dataset with global abundance dataset of 9,700 bird species. Even if taxonomic order were used as a lower-level classification standard, the original LPI weighting process would still be unachievable for my work due to the difficulty in obtaining the number of species from different orders within each realm. Secondly, while previous LPI studies filtered population abundance sequences to represent global populations, the resulting datasets still largely reflected abundance changes across most species within each class. In order to get robust data for calculating species global abundance time sequences, my dataset was reduced to only 304 species after multiple filtering steps. Data of selected 304 species is insufficient to represent overall biomass changes in global scale but could serves as more robust data foundation when calculating weighted LPI by abundance according to the hypothesis about abundance's influence on species' growth. Moreover, the weighted LPI by abundance in my research resolved the problem that there is no comparability between abundance time sequences of different species.

4.4 Adjustment for Hill number

In the process of estimating global abundance sequences, the time when the global abundance was selected is without the time range of species lambda dataset, so there is bias between the global abundance sequences I obtained and the actual abundance data. Although the species abundance dataset derived from data scaling may not strictly serves as perfect resource for original Hill number analysis, the trend of abundance sequences and abundance's magnitude are correct because I value sequences are results of cumulative calculation, which could simulate the growth of species abundance. On the other hand, the calculation process for the Hill number in this research does account for relative abundance. Relative abundance can serve as a measure of species' influence (James & Rathbun 1981), which can be adjusted by modifying the q values in the Hill number calculation (Chao et al. 2014). Therefore, although it is not the original Hill number, it remains a reliable metric for reflecting changes in species biodiversity.

4.5 Living Planet Indexes

In all species group, dominant species group and rare species group, there are remarkable differences between the original LPI and weighted LPI by abundance. There are more negative values in growth rate sequences of weighted LPIs than original ones in these mentioned groups. These results demonstrate that in groups with lower species evenness level or sensitivity, the weighting process by abundance could make significant difference in LPI values.

Previous research shows that when a community is primarily composed of dominant species, the level of evenness tends to stabilize because the influence of rare species on evenness is minimal and the dominant species' abundance is stable (Chao & Ricotta 2019). In such cases, the development of the community is reflected on changes in other factors of biodiversity such as biomass rather than species diversity. As a result, trends of both original and abundance-weighted LPIs differ significantly from the Hill number trend in dominant species groups. Research indicates that genetic diversity and species diversity interact to affect biomass (Crawford & Rudgers 2012). As a result, the weighted LPI by abundance should be used to track biodiversity changes over time in communities with dominant species, as it encompasses multiple aspects of biodiversity, including genetic diversity.

4.6 Hill numbers

In my research, species richness within each abundance-based species group remained stable at most time points. Changes in the Hill number were primarily driven by fluctuations in species evenness, as no extinctions were observed according to the species lambda dataset. Consequently, as the q value increases, the fluctuations in sequences become more pronounced, varying in magnitude across different abundance-based species groups due to the increased weighting of evenness and the differing sensitivities of these groups to evenness. The Hill number sequences across different abundance-based groups do not follow the inverse trend of the sequence of standard deviations of the abundance distribution. It means that the trend of species evenness indicated by Hill number is different from the trend shown by abundance distribution. Changes in Hill number sequences are particularly acute in the middle abundance and less rare species groups compared to others. Thus, it is the combined effect of both abundance distribution and other factors including evenness sensitivity that leads to the varying growth rates of Hill number sequences across different abundance-based species groups at different time points.

4.7 Pielou evenness and Ip(λ)

Pielou's evenness is typically considered a reliable index of evenness because it is less affected by species richness (Gosselin 2006). However, in my research, its changing rates consistently remained low, indicating that the stability in Pielou's evenness may be largely influenced by the unchanged species richness. Therefore, I did not pursue further analysis on it.

Generalized levels of relative evenness expressed by Ip(λ) in different abundance-based species groups are similar to different groups' Pielou evenness levels while development trends of Ip(λ) indexes are similar to those of Hill numbers(q=2). So my research illustrates that Ip(λ) could indicate relative evenness to large extend.

4.8 Evenness Sensitivity Analysis

In my research, I measured the sensitivity of different abundance-based species groups to evenness by analyzing how the growth rates of Hill number sequences changed as the q value increased. The reason why I chose this measurement is that these changes in growth rates were more significant than those observed in Ip(λ) sequences when the λ value varied and changes in Hill numbers while q value changes are not effected by species richness due to its unchanged level.

According to the GLM fitting results, P-values of coefficient for abundance distribution could reach significant level within abundance-based species groups that have lower standard deviations in abundance distribution. This suggests that though the evenness measurements typically consider other aspects of species diversity (Blowes et al. 2022), in cases where abundance distribution is highly even and could cause extremely more significant effect on relative evenness than other coefficients, it is possible to represent species evenness by abundance distribution. Because in that case, abundance distribution is the most dominant factor of relative evenness. The p-values of both coefficients are less than 0.05

in all species group and rare species group, not reaching significant level. It shows that in these cases, species evenness has weak connections with both abundance distribution and sensitivity to evenness. So species evenness do have connections with other factors of biodiversity in ecological aspect Moreover, further research on evenness sensitivity is needed to explore additional dimensions of species evenness as coefficients of evenness sensitivity are not significant in most abundance-based species groups except for rare species group.

4.9 Broader Implications and Future Directions

Species diversity is a priority in biodiversity policy across many contexts (Henle et al. 2013). Evenness sensitivity may link other biodiversity factors to species evenness and serve as a valuable indicator for biodiversity policymakers, as it helps identify the primary drivers of species evenness to some extent. If the coefficient of evenness sensitivity is much more significant than that of abundance distribution in the GLM, the community's relative evenness is mainly influenced by evenness sensitivity, then efforts to adjust evenness levels should focus on other aspects of the community's biodiversity to adjust community's sensitivity to evenness rather than its abundance distribution. More robust models need to be developed in the future to analyze different dimensions of species evenness, allowing the fitting results to identify the main factors of evenness through p-values. However, the methodology for calculating evenness sensitivity should be adapted to the community's status. In cases where species richness fluctuates significantly, the growth rate changes of Ip(λ) sequences when λ values change should be referred to as a factor of community's sensitivity to evenness because the calculation of Ip(λ) is independent from species richness. Additionally, further analysis on more dimensions of species evenness is needed to develop a robust indicator of evenness sensitivity, as its p-values do not reach the significant level in several abundance-based species groups in my research.

5 Data and Code Availability

The Living Planet Dataset (LPD), foundational for all analyses, can be downloaded from `https://www.livingplanetindex.org`. The global species lambda dataset, derived from the LPD, is accessible via the `rlpi` R package at `https://github.com/Zoological-Society-of-London/rlpi`. The dataset on birds' body mass can be downloaded from `https://academic.oup.com/biolinnean/article/163/data`. The LPI completeness data, indicating the percentage of each population in the LPD that represents the global populations of the corresponding species, is yet to be formally published. The results of the datasets collated in this research can be found in this repositories: `https://github.com/Zhin-Xu/Fanshu_Research_Project2`.

References

Andelman, S. J. & Willig, M. R. (2003), 'Present patterns and future prospects for biodiversity in the western hemisphere', *Ecology letters* **6**(9), 818–824.

Blowes, S. A., Daskalova, G. N., Dornelas, M., Engel, T., Gotelli, N. J., Magurran, A. E., Martins, I. S., McGill, B., McGlinn, D. J., Sagouis, A. et al. (2022), 'Local biodiversity change reflects interactions among changing abundance, evenness, and richness', *Ecology* **103**(12), e3820.

Bock, C. E., Jones, Z. F. & Bock, J. H. (2007), 'Relationships between species richness, evenness, and abundance in a southwestern savanna', *Ecology* **88**(5), 1322–1327.

Bromiley, P., Thacker, N. & Bouhova-Thacker, E. (2004), 'Shannon entropy, renyi entropy, and information', *Statistics and Inf. Series (2004-004)* **9**(2004), 2–8.

Brown, J. H., Gillooly, J. F., Allen, A. P., Savage, V. M. & West, G. B. (2004), 'Toward a metabolic theory of ecology', *Ecology* **85**(7), 1771–1789.

Buchanan, G. M., Butchart, S. H., Chandler, G. & Gregory, R. D. (2020), 'Assessment of national-level progress towards elements of the aichi biodiversity targets', *Ecological Indicators* **116**, 106497.

Buschke, F., Hagan, J., Santini, L. et al. (2021), 'Random population fluctuations bias the living planet index. nat ecol evol 5: 1145–1152'.

Butchart, S. H., Walpole, M., Collen, B., Van Strien, A., Scharlemann, J. P., Almond, R. E., Baillie, J. E., Bomhard, B., Brown, C., Bruno, J. et al. (2010), 'Global biodiversity: indicators of recent declines', *Science* **328**(5982), 1164–1168.

Callaghan, C. T., Nakagawa, S. & Cornwell, W. K. (2021), 'Global abundance estimates for 9,700 bird species', *Proceedings of the National Academy of Sciences* **118**(21), e2023170118.

CBD, U. (2010), Strategic plan for biodiversity 2011–2020 and the aichi targets, *in* 'Report of the Tenth Meeting of the Conference of the Parties to the Convention on Biological Diversity'.

Chandra, A. & Idrisova, A. (2011), 'Convention on biological diversity: a review of national challenges and opportunities for implementation', *Biodiversity and Conservation* **20**, 3295–3316.

Chao, A., Chiu, C.-H. & Jost, L. (2014), 'Unifying species diversity, phylogenetic diversity, functional diversity, and related similarity and differentiation measures through hill numbers', *Annual review of ecology, evolution, and systematics* **45**(1), 297–324.

Chao, A. & Ricotta, C. (2019), 'Quantifying evenness and linking it to diversity, beta diversity, and similarity', *Ecology* **100**(12), e02852.

Chaytor, B., Gerster, R. & Herzog, T. (2002), 'The convention on biological diversity: Exploring the creation of a mediation mechanism', *J. World Intell. Prop.* **5**, 157.

Collen, B., Loh, J., Whitmee, S., McRAE, L., Amin, R. & Baillie, J. E. (2009), 'Monitoring change in vertebrate abundance: the living planet index', *Conservation Biology* **23**(2), 317–327.

Crawford, K. M. & Rudgers, J. A. (2012), 'Plant species diversity and genetic diversity within a dominant species interactively affect plant community biomass', *Journal of Ecology* **100**(6), 1512–1521.

Cressie, N. & Read, T. R. (1984), 'Multinomial goodness-of-fit tests', *Journal of the Royal Statistical Society Series B: Statistical Methodology* **46**(3), 440–464.

Gosselin, F. (2006), 'An assessment of the dependence of evenness indices on species richness', *Journal of theoretical biology* **242**(3), 591–597.

Hallmann, C. A., Ssymank, A., Sorg, M., de Kroon, H. & Jongejans, E. (2021), 'Insect biomass decline scaled to species diversity: General patterns derived from a hoverfly community', *Proceedings of the National Academy of Sciences* **118**(2), e2002554117.

Henle, K., Bauch, B., Auliya, M., Külvik, M., Pe'er, G., Schmeller, D. S. & Framstad, E. (2013), 'Priorities for biodiversity monitoring in europe: A review of supranational policies and a novel scheme for integrative prioritization', *Ecological Indicators* **33**, 5–18.

Hubbell, S. P. (2001), 'The unified neutral theory of biodiversity and biogeography princeton'.

James, F. C. & Rathbun, S. (1981), 'Rarefaction, relative abundance, and diversity of avian communities', *The Auk* **98**(4), 785–800.

Joly, C. A. (2022), 'The kunming-montréal global biodiversity framework'.

Jost, L. (2006), 'Entropy and diversity', *Oikos* **113**(2), 363–375.

Jost, L. (2010), 'The relation between evenness and diversity', *Diversity* **2**(2), 207–232.

Keylock, C. (2005), 'Simpson diversity and the shannon–wiener index as special cases of a generalized entropy', *Oikos* **109**(1), 203–207.

Loh, J., Green, R. E., Ricketts, T., Lamoreux, J., Jenkins, M., Kapos, V. & Randers, J. (2005), 'The living planet index: using species population time series to track trends in biodiversity', *Philosophical Transactions of the Royal Society B: Biological Sciences* **360**(1454), 289–295.

MacArthur, R. H. (1965), 'Patterns of species diversity', *Biological reviews* **40**(4), 510–533.

McRae, L., Deinet, S. & Freeman, R. (2017), 'The diversity-weighted living planet index: controlling for taxonomic bias in a global biodiversity indicator', *PloS one* **12**(1), e0169156.

Michonneau, F., Brown, J. W. & Winter, D. J. (2016), 'rotl: an r package to interact with the open tree of life data', *Methods in Ecology and Evolution* **7**(12), 1476–1481.

Morales-Castilla, I., Rodriguez, M. A. & Hawkins, B. A. (2012), 'Deep phylogeny, net primary productivity, and global body size gradient in birds', *Biological Journal of the Linnean Society* **106**(4), 880–892.

Mouillot, D. & Wilson, J. B. (2002), ‘Can we tell how a community was constructed? a comparison of five evenness indices for their ability to identify theoretical models of community construction’, *Theoretical Population Biology* **61**(2), 141–151.

Purvis, A. & Hector, A. (2000), ‘Getting the measure of biodiversity’, *Nature* **405**(6783), 212–219.

R Core Team (2023), *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria. **URL:** *https://www.R-project.org/*

Read, T. R. & Cressie, N. A. (2012), *Goodness-of-fit statistics for discrete multivariate data*, Springer Science & Business Media.

Rosendal, G. K. (2000), *The convention on biological diversity and developing countries*, Vol. 25, Springer Science & Business Media.

Sarkar, S. (2002), ‘Defining” biodiversity”; assessing biodiversity’, *The Monist* **85**(1), 131–155.

Studeny, A., Buckland, S., Illian, J., Johnston, A. & Magurran, A. (2011), ‘Goodness-of-fit measures of evenness: a new tool for exploring changes in community structure’, *Ecosphere* **2**(2), 1–19.

Tuomisto, H. (2010), ‘A consistent terminology for quantifying species diversity? yes, it does exist’, *Oecologia* **164**(4), 853–860.

Wilsey, B. J., Chalcraft, D. R., Bowles, C. M. & Willig, M. R. (2005), ‘Relationships among indices suggest that richness is an incomplete surrogate for grassland biodiversity’, *Ecology* **86**(5), 1178–1184.

Wilsey, B. J. & Potvin, C. (2000), ‘Biodiversity and ecosystem functioning: importance of species evenness in an old field’, *Ecology* **81**(4), 887–892.

Wilson, E. (1995), ‘Peter fm, 1988. biodiversity’, *National Academic Press, Washington. J. Ahern/Landscape and Urban Planning* **33**, 131–155.