

A Proposed Framework for Integrating Productivity and Niche-Based Measures in Functional Diversity Analysis: Simulation-Based Characterization

Aryan Ayyanger

aryan.ayyanger@gmail.com

Independent Scholar

Ahdav Vijay

Independent Scholar

Research Article

Keywords:

Posted Date: September 23rd, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-10821215/v2>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Additional Declarations: The authors declare no competing interests.

A Proposed Framework for Integrating Productivity and Niche-Based Measures in Functional Diversity Analysis: Simulation-Based Characterization

Aryan Ayyanger¹, Ahdav Vijay²

Abstract

Functional diversity indices provide quantitative approaches for describing ecological differences among organisms, but many existing frameworks rely on predefined traits or abundance distributions. This study proposes a Functional Diversity Index (FDI) that integrates oxygen consumption and dietary niche composition within Rao's Quadratic Entropy structure weighted by Pianka-based niche dissimilarity. The proposed index was examined using a NetLogo simulation designed to generate randomized distributions of organisms, oxygen consumption, and dietary resource use. Across 91 randomized distributions, FDI values ranged from 0.1807 to 0.3151. Regression analysis explained 84.13% of the variation in simulated FDI values (adjusted $R^2 = 81.5\%$) demonstrating strong internal consistency, while Classification and Regression Tree (CART) analysis identified threshold-like transitions associated with dietary composition within the randomized ecosystem. Additionally, K-means clustering was used to explore patterns among simulated ecological variables, and an optimization procedure was developed to prove the concept of the FDI. These analyses demonstrate how the proposed FDI responds to changes in consumption distributions under the specific simulation conditions. Because the variables used to construct the FDI are also present among the simulated predictors, the present analysis characterized the mathematical behavior of a novel index rather than independently establishing its ecological validity. The framework therefore provides a preliminary model for integrating energy-related consumption and dietary niche structure into functional diversity analysis.

Introduction

The present world is marked by exponential technological growth, and one major consequence of these technological advancements is the rapid environmental changes occurring across the planet. These shifts are driving deep structural changes in our ecosystems. We are beginning to see system-level shifts in productivity, resource availability, and species interactions. Ecologists have emphasized that older methods of simply counting species, without examining how each species uses resources, processes energy, and interacts functionally, are no longer sufficient. There is an immediate need for a precise way to measure changes in biodiversity that can drive meaningful actions to preserve ecosystem stability—an index that can take advantage of the current growth in data science and machine learning.

A major challenge is that functional roles in our ecosystems are extremely dynamic. Resource use shifts with environmental conditions, metabolic demands vary across individuals, and niche overlap can expand or contract depending on productivity levels. Traditional trait-based frameworks, including entropy-based measures such as Rao's Quadratic Entropy, provide valuable structure but depend heavily on predetermined trait matrices that do not adjust when organisms alter their consumption patterns. Similarly, niche-overlap models like Pianka's Index quantify shared resource use but are seldom linked to energy-flow variables such as Gross Primary Productivity (GPP), even though metabolic theory shows that energy processing is fundamental to ecological organization.

Current functional diversity metrics struggle to represent ecosystems with flexible resource consumption, variable energy flow, and functional distance shifts as environmental conditions change. As a result, they often fail to detect the environmental transitional threshold and nonlinear responses that characterize many ecological systems. These limitations are especially important in contexts where productivity and resource partitioning—not species richness alone—govern resilience.

The motivation for this study arose from this gap. Instead of frameworks developed on static, empirical traits and taxonomic richness, this study works to model continuous, dynamic scenarios within an agent-based architecture. The goal was to create a diversity index capable of capturing how ecosystems reorganize when resource use shifts, even when species identities remain the same. By integrating oxygen-based productivity measures with multi-resource dietary niches, the study attempts to develop a functional index that is sensitive to ecological shifts and clearly identifies tipping points and plastic transformations.

This perspective guided the development of our Functional Diversity Index (FDI), which combines an evenness term derived from oxygen consumption with a quadratic-entropy structure weighted by Pianka-based niche dissimilarity. The index was designed to evaluate functional structure under conditions where organisms differ not only in abundance, but also in how they partition resources and contribute to overall energy flow. Simulation-based testing and statistical analysis helped examine whether such an index could identify functional shifts and reveal the underlying drivers of ecosystem organization.

Literature Review

Limitations of Traditional, Trait-Based Functional Diversity Metrics

Traditional biodiversity metrics primarily focus on species abundance as a descriptive measure of ecological diversity rather than the ecological roles these organisms possess. Many of the most established indices and models, such as species richness and Shannon diversity, quantify the mathematical distributions of these ecosystems in the context of evenness, but they do not capture how organisms function within their environment through distinct roles.

Rao's Quadratic Entropy, introduced by C.R. Rao (1982), measures diversity using pairwise functional differences between species weighted by abundant counts (3). Earlier work by Laszlo Szeidl and Carlo Ricotta (2009) found that traditional indices have shortcomings when representing ecological variation and that entropy-based frameworks like Rao's better represent functional variation by incorporating dissimilarity (12). However, the results of this study also show sensitivity in these indices to the definitions of ecological distance, meaning results dependent on predefined trait structures.

Botta-Dukát (2005) tested Rao's index across multiple simulated communities and found that the model's sensitivity to both trait differences and species abundance satisfied the core criteria for functional diversity (1). However, he also found that Rao's Quadratic Entropy has diminishing

interpretive success when many correlated traits are included, since redundancy falsely magnifies rather than clarifying functional separation for ecological representation.

Many works extend Rao's applications beyond ecological representation into external scientific analysis. For example, Smouse (2017) expanded the framework for analyzing genetic diversity where Rao's index had success in capturing the hierarchical variation across species levels (16). However, the study also noted that the model depends strongly on defined matrices, which limits the flexibility in more dynamic ecosystems. This indicates a key limitation in the literature: although trait-based diversity indices elevate ecological representations, they remain constrained by static assumptions rather than modeling consistently changing ecosystems with highest accuracy.

Niche Overlap and Species Interactions

Another approach to understanding biodiversity is based on species interactions, shared resource usage, and ecological similarity. The Pianka's Index has been traditionally used for quantifying the niche overlap present in an ecosystem by comparing the amount of resources used by different species (11). This helped provide a mathematical approach for observing the extent of competition between species within a community. Furthermore, MacArthur and Levins found that coexisting species are strongly shaped by the amount of overlap in resource use, especially concerning ecological niches (9).

Other work has found that niche overlap is not a static characteristic. Fattorini and other researchers found that niche overlap can shift significantly, especially after environmental disturbances, which can change species interactions and the ecological makeup of a community (5). Additionally, Schoener (1974) established that niche overlap is context-dependent and can differ across many environments based on the availability of resources and other constraints (15).

Considering all of this together, these findings indicate that niche overlap shouldn't be perceived as a static aspect of species pairs. Instead, they should be viewed as a dynamic result of differing environmental conditions and the activities within communities. Therefore, biodiversity doesn't only result from species and their traits but also actively changing networks in the ecosystem.

Energy Flow, Productivity (GPP)

Optimizing ecological efficiency and relationships not only focuses on traits and niche overlap between species, but also requires analysis on energy-processing systems that constrain ecological flow and productivity. Ecosystems can be modeled through these energy-based frames though Gross Primary Productivity, metabolic rates, and respiration where limited resources satisfy ecological demands.

Odum (1969) also established that ecological stability is influenced by energy flow pathways: he argued that succession and ecosystem maturity are influenced by energy availability and respiration balances (10). His work demonstrated that developed ecological systems consistently have greater stability in energy efficient scenarios rather than purely high species richness states.

Another study by Lindeman (1942) quantified the efficiencies in energy transfer between trophic levels, and what ideal ecological distributions are for maintaining sustainability (8). His findings showed how productivity losses constrain food web structures, so maintaining plant biomass is crucial for sustenance. Scientists have worked to measure these energy changes and interactions through dynamic and functional manners. Similarly to the research done by Lindeman (1942), Running et al. (2004) models energy through measurable ecological parameters (13). Specifically, the study utilized satellite-derived productivity to demonstrate that GPP has the capacity to model activity and carbon exchange. Due to the relationships between oxygen exchange and photosynthetic productivity, these findings support the use of oxygen-based variables in ecological models.

The relationship between energy-processing and ecological organization was also expanded by the Metabolic Theory of Ecology (2). Their study supported that metabolic rates scale in predictable manners with temperature across organisms and body size. These organism specific energy transfers allow ecological processes such as growth, respiration, and consumption to be mathematically modeled. The authors found that metabolic scaling laws influencing ecological functions is strongly relevant, and that this perspective allows for dynamic ecosystem representation.

Nonlinear Ecosystem Dynamics

Beyond productivity itself, dynamic ecosystem representation involves analysis on the nonlinear responses that ecological systems have to environmental disturbances and functional imbalances. Holling (1973) introduced resilience theory and demonstrated that ecosystems remain stable across a range of environmental conditions before rapidly shifting into alternative stable states once critical thresholds are exceeded (6). This work posited a novel perspective on previous ecological assumptions by proposing that resilience depends on a system's ability to absorb disturbances with functional collapse.

Scheffer et al. (2001) further expanded this concept by focusing on catastrophic regime shifts in defined ecosystems such as lakes and coral reefs (14). The study's findings suggested that small environmental pressures may trigger disproportionately large structural changes once tipping points are exceeded. This suggests that ecological transitions are frequently abrupt rather than gradual, making threshold-sensitive models essential for monitoring purposes. Furthermore, Dakos et al. (2012) formally identified measurable warning signals preceding these ecological tipping points: these included increased variance, stronger correlation patterns, and compromised recovery rates (4). This study's efforts to quantify these thresholds in an analytical system suggest that nonlinear instabilities may be detectable before complete ecological collapse occurs.

Collectively, these studies indicate that ecosystem stability is not purely dependent on biodiversity itself, but also resilience to threshold-sensitive interactions. However, many traditional biodiversity indices and models primarily measure static abundance distributions while overlooking the nonlinear ecological transitions present in realistic environments. Consequently, integrating a productivity-derived perspective with dynamic resource-consumption tipping points may provide a more responsive framework for representing instability and thus assists in ecological optimization.

Current Gaps

Traditionally, biodiversity is measured through abundance-based indices such as species richness, Shannon's Diversity Index, and others, which focus on finding how many species exist and how evenly they are distributed. However, a disadvantage of these indices is that they often miss functional roles a species has within an environment – roles that elucidate efficiency, fluidity, and overall health of the biosphere. These functional roles describe how species contribute to energy flow and the environment (through activities like pollination or seed dispersal).

Currently, advanced trait metrics like Rao's Quadratic Entropy and Pianka's Index pursue incorporating a dissimilarity factor and niche overlap. Although this does improve the representation of the functions within a species, current models primarily treat traits as static and fail to capture multi-resource consumption spanning multiple trophic levels.

These limitations pose multiple gaps in measuring functional diversity. There is a lack of rigorous, clear, mathematically based indices that numerically model functional diversity based on energy and resource consumption. Additionally, most indices also exclude GPP (Gross Primary Productivity), which is derived from oxygen or energy flow variables. Due to these factors, current models struggle to detect threshold-driven shifts within an ecosystem's functionality and may fail to fully understand the ecosystem's progression.

Ultimately, our goal was to develop a Functional Diversity Index which incorporates multi resources consumption and oxygen consumption rates, which are derived from GPP. The model combines Rao's Quadratic Entropy (functional dissimilarity weighted by abundance) along with Pianka's Index (niche overlap) to construct a testable framework that can model functional shifts in a simulated terrestrial environment. This leads to the research question: How would a Functional Diversity Index based on multi-resource consumption and oxygen consumption rates derived from GPP identify threshold-driven shifts in ecosystem functional structure within a simulated terrestrial ecosystem?

Functional Diversity Index

The Functional Diversity Index was an integrated model of multiple ecological aspects to cover the diversity of an ecosystem through both interspecies and intraspecies analysis. The first summation term is a measure of functional evenness based on the proportion of oxygen consumed by a species from all of the photosynthetic production by a simulated plant biomass. Instead of focusing on the raw abundance of a species as a proportion, the oxygen consumption evenness component sets the base for the ecosystem's functional biodiversity. As shown in Figure 1, each of the proportions p_i for ($i=1..n$) is squared and subtracted from 1, so a higher value for the term occurs when the sum of squared proportions are minimized.

The second term of the index was created with quadratic entropy structure and Pianka's Index for the dissimilarity value. As shown by Figure 1, each term in the summation is the product between the consumption proportions of two distinct organisms and a dissimilarity value based

on the organisms' dietary niches. The proportions used are the relative oxygen consumption for both the first and second organism, which is defined from the randomization process in the simulation. It's worth noting the multiplier at the front being by one-half mathematically avoids overcounting when distinguishing between permutations and combinations.

$$\frac{1}{2} \left(1 - \sum_{i=1}^n p_i^2 \right) \cdot \left(\sum_{i=1}^n \sum_{j=1}^n p_i \cdot p_j \cdot d_{ij} \right)$$

Figure 1. Novel functional diversity index: combines evenness, Rao's Quadratic Entropy, and Pianka's Index

The d_{ij} value determined by the dietary composition of the two distinct organisms is calculated by the use of Pianka's index, as shown in Figure 2. This niche value is calculated by first determining the niche similarity between two species (o_{ij}), and then subtracting the value from 1 to measure dietary diversity. This calculated term is repeated for all pairs of organisms. It is worth noting that the niche similarity between the same organisms is exactly 1, so the term d_{xx} is equal to 0. Additionally, the $\frac{1}{2}$ multiplier at the front is to avoid overcounting between the cases d_{ab} and d_{ba} where the niche values are identical and counted twice.

$$o_{ij} = \frac{\left(\sum_{i=1}^n p_{ik} \cdot p_{jk} \right)}{\left(\sqrt{\left(\sum_{i=1}^n p_{ik}^2 \right) \left(\sum_{j=1}^n p_{jk}^2 \right)} \right)}$$

Figure 2. Pianka's Index for measuring niche diversity between two organisms. It is calculated based on dietary consumption distributions.

Methodology

The derived functional diversity index was implemented in a NetLogo-based simulation to characterize the behavior of the proposed mathematical framework under controlled conditions. The index integrates an evenness term based on oxygen consumption, a quadratic entropy-based structure, and Pianka's niche dissimilarity for dietary compositions. The framework was designed to accommodate variation in species richness and resource types and was evaluated using simulated variables constrained by the parameters of the model. The one-half multiplier in the quadratic entropy component prevents double-counting when distinguishing between pairwise permutations and combinations. Together, these components provide a mathematical framework for examining functional variation across simulated ecosystem distributions.

The simulation development began with defining the fundamental variables and parameters required for randomized ecosystem generation. These attributes, developed in version 1.0, included oxygen output, established user control, and defined important variables, including Net

Primary Productivity (NPP) and Gross Primary Productivity (GPP). In version 1.1, distinct organisms were visually represented with unique sizes and colors for a more interactive user-defined interface. Later updates incorporated user-input variables including sample size of organisms from 2 to 1000, and multiple resource types with randomized distributions. The final simulation generated output variables including NPP, GPP, average-consumption, and the Functional Diversity Index value, all of which are displayed in the final interface.

```
There are 100 Species-1 organisms
There are 109 Species-2 organisms
There are 91 Species-3 organisms

The GPP of this ecosystem is 30689.39466471508 units
The NPP of this ecosystem is 14613.997459388132 units

Average Species-1 oxygen consumption: 14.736656172063604
Average Species-1 food-1 consumption: 0.6019667323891991
Average Species-1 food-2 consumption: 0.13125882038087514
Average Species-1 food-3 consumption: 0.2667744472299258

Average Species-2 oxygen consumption: 48.753895264597524
Average Species-2 food-1 consumption: 0.13467757942547628
Average Species-2 food-2 consumption: 0.6132548296126323
Average Species-2 food-3 consumption: 0.2520675909618918

Average Species-3 oxygen consumption: 86.001728113633330
Average Species-3 food-1 consumption: 0.2634990297266635
Average Species-3 food-2 consumption: 0.1327007493730494
Average Species-3 food-3 consumption: 0.6038002209002868

Evenness Factor = 0.9953999629606903
Dissimilarity Factor = 0.1478932650011753
Functional Diversity Index Value = 0.14721295050430547
Shannon Diversity Index Value = 1.0959086318068039
```

Figure 3. Sample output of the simulation software. Contains variable types, species amounts, and output values.

Procedure

The algorithmic workflow integrates user-defined population parameters with automated stochastic assignments. First, the user entered the number of organisms into the simulation interface. The program then would conduct the random assignment procedures: randomly assigning oxygen and dietary compositions to each of the distinct organisms. This randomization generated varied distributions without assigning organisms to predetermined combinations or traits, allowing characteristics to be classified after assignment. These classifications were in three main categories: Species-1, Species-2, and Species-3, each with defined ranges across the modeled variables. For example, Species-1 consumed between 0 and 30 units of oxygen, whereas Species-2 consumed between 30 and 70 units, while allowing variation within each category. After these preliminary steps, the algorithm calculated the evenness term based on the oxygen proportions as a guiding value and the pairwise quadratic entropy component using Pianka's dissimilarity value. The simulation also calculated the Shannon diversity index for each specific ecosystem distribution for the purposes of comparing with a conventional diversity measure. This process occurred 91 times for distinct distributions where the statistical values would be derived.

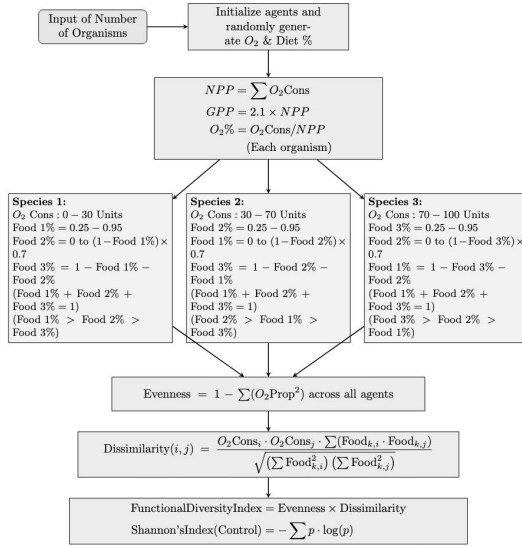


Figure 4. Algorithmic framework for simulation creation. Includes crucial steps and the species classification: Species 1, Species 2, and Species 3.

The primary analysis focused on the resulting values of the Functional Diversity Index and their variation across simulated ecological distributions. Within the wide varieties of randomized distributions, two hypothetical cases established reference conditions for interpreting the randomized results. The first was a baseline scenario defined by a singular organism per species, fixed oxygen and dietary constraints, and isolated calculation of the index. This case represented maximum differentiation under equal representation, which was a core simulation standard for future analysis. The second scenario represented the maximum FDI value obtainable under the practical constraints of the simulation. Although this value was not a formal theoretical maximum, the value represented the maximum under the model's specified restrictions. In this scenario, 1000 organisms were randomized across specific species clusters, with populations generally exhibiting greater diversity. These two hypothetical scenarios served as a reference for examining the range of FDI values and central to subsequent optimization analysis.

Due to the multivariable composition of the functional diversity index and the various degrees of influence amongst categories of a specific variable, regression analysis was used to characterize relationships between the simulated variables and the FDI. Species-based variables, including oxygen consumption and dietary composition, as well as broader ecological parameters were grouped according to their statistical significance across three categories: statistically significant predictors, suggestive near-significant predictors, and non-significant variables. These classifications provided insight into how the index responded to variation within the simulated ecosystem. However because oxygen consumption and dietary composition are incorporated directly into the construction of FDI, their statistical relationships cannot independently establish ecological validity. Instead regression was used to characterize the sensitivity of the proposed index to simulated variables. Several of the variables exhibited high VIF values, indicating multicollinearity and limiting the interpretation of individual relationships. Due to severe multicollinearity among abundance metrics (elevated VIF > 1000), non-parametric Classification and Regression Tree (CART) analysis was conducted to isolate nonlinear patterns demarcate functional regime shifts. The resulting trinodal identified points at which simulated

functional-diversity regimes changed under different combinations of resource-use variables. These thresholds characterize behavior within the model and are not interpreted as empirically established ecological tipping points. Relative variable importance of these distinct variables was compared in standardized graphs for enhanced comparison.

Although regression, CART, and FDI analyses provided information about relationships among the defined variables, they did not fully capture multidimensional patterns across changing ecosystem distributions. For example, the regression analysis considered the importance of the pre-determined variables and their classification among the species categories, including food and oxygen consumptions. To examine greater multidimensional complexity, a K-means clustering approach was integrated as an exploratory method for organizing simulated distribution according to similarities across variables. K-means clustering organizes the data points into distinct variable clusters, indicating repeating patterns or trends that may represent underlying relationships among dimensions. This approach allowed for multidimensional representation by using integrated Euclidean distances to update the center point; thus, the approach was critical for the procedure's practical use across distinct distributions since the variables can be ecologically specific. Within functional ecology, K-means clustering provides a rigorous method for partitioning multi-dimensional trait spaces and organizing species into functional groups (7).

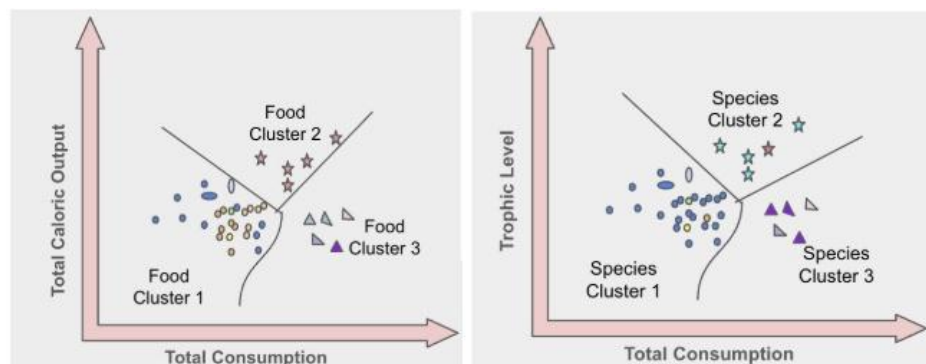


Figure 5. K-means clustering approach for two different dimensions, being the Total Caloric Output and the Trophic Level. The data in these graphs are organized into three clusters.

The preceding analyses provided a proof-of-concept optimization procedure using oxygen consumption rates. Rather than establishing an empirically validated sustainability intervention, the optimization model detailed how the proposed framework could respond to simulated changes in species dominance. The model used a recursive procedure involving the two most dominant species within a given cycle. After calculating the ratio between the most dominant to the second-most-dominant, the algorithm adjusted the dominant proportion according to mathematically defined conditions. The casework tree considered three possible adaptations: maintaining the dominant species-proportion, applying a smoothing function, or multiplying by 0.75 to produce a less abrupt change. The smoothing function itself was incorporated to allow continuous species regulation and was determined using established parameters in the model. Specific conditions and guidelines governing this procedure are presented in **Figure 6**. This optimization therefore demonstrates a potential computational application of the FDI framework

within a simulated ecosystem rather than establishing that the procedure would increase sustainability in natural ecosystems.

Definitions

1. Define P_1 as the proportion of holistic consumption of the species with the highest oxygen consumption
2. Define P_2 as the proportion of the holistic consumption of the species with the second highest oxygen consumption
3. Define r = dominance ratio as P_1/P_2
4. Smoothing function: component of multiplier for P_1 : $p(r) = 1/(1 + 5(r - 1)^2)$
5. Final modifier value $a = 0.3$ is adjustment index for smooth calculation

Algorithm The model will calculate a gradual optimization value for the larger population and will choose the minimum between the generated value, $0.75 * P_1$, and the original value.

Note: The value of $0.75 * P_1$ will only be considered when $0.75P_1 > P_2$

Final Function: Final optimization multiplier: $f(r) = 1 - 0.3 + 0.3 * (p(r))$

Formula : $f(r) = 0.7 + 0.3 * (1/(1 + 5(r - 1)^2))$

Approach

Three values will be considered as the possible final optimization value for the highest population P_1 : $0.75 * P_1$, $P_1 * f(r)$, and P_1 .

Process

1. The value $0.75 * P_1$ and $P_1 * f(r)$ will only be considered if it is greater than P_2
2. $\min(0.75 * P_1, f(r) * P_1, P_1)$ will be selected as the new value
3. This will be repeated recursively for all the populations by shifting the higher population to the next highest: the analysis will occur for P_2 and P_3 , P_3 and $P_4, \dots, P_{n-1}$ and P_n for all n where n is the species richness
4. This recursive analysis occurs for the first few dominant populations: can range from 1 to 3 depending on ecosystem distribution

Reasoning

1. The value of $a = 0.3$ represents the minimum multiplier. The value is chosen for moderate changes in the ecosystem that avoid abrupt ecosystem adaptations
2. The value of 5 in the denominator of the smoothing function allows for more rapid changes to occur at moderate values of r . This is important for the nature of the multiplier: drastically decrease higher ratios of r while preserving lower values

Figure 6. Definitions, algorithm, approach, process, and reasoning for the optimization model. Contains the smoothing function, important parameters, and the tested procedure.

Results

In our simulation results, there were 91 randomized ecosystem distributions, which allowed us to examine the behavior of the Functional Diversity Index (FDI) under multiple different combinations of the oxygen consumption rate, species abundance, and the composition of the diet.

Baseline and Maximum Scenarios

In the base scenario, the model initialized one organism for each species cluster in order to isolate the differences in functionality without the influence of the population structure. Under the scenarios with minimal conditions, the ecosystem produced an FDI value of 0.1807 (**Figure 7**). This value represents the lower range of FDI values observed within the simulated conditions, where species are equally represented, and resource consumption patterns are relatively constrained. Under the maximum scenario with 1,000 organisms distributed across three species clusters with random trait assignments, the model gave an FDI value of 0.3151, which showed a greater functional differentiation under the larger population distributions (**Figure 7**). This increase from the baseline also shows that the FDI increases under the greater organism representation and randomized resource-use configurations used in the maximum scenario. Altogether, these scenarios helped create an estimated operational range of 0.1807 to 0.3151 for the FDI, which helps provide a reference for later interpretation.

Functional Diversity Index		INPUT	Functional Diversity Index		INPUT
# of Species 1	1		# of Species 1	334	
# of Species 2	1		# of Species 2	331	
# of Species 3	1		# of Species 3	117	
Sp1 O2 Consumption	25		Sp1 O2 Consumption	25	
Sp1 Fd1 Consumption	0.95		Sp1 Fd1 Consumption	0.95	
Sp1 Fa2 Consumption	0.035		Sp1 Fa2 Consumption	0.035	
Sp1 Fd3 Consumption	0.015		Sp1 Fd3 Consumption	0.015	
Sp2 O2 Consumption	63		Sp2 O2 Consumption	26	
Sp2 Fd1 Consumption	0.035		Sp2 Fd1 Consumption	0.035	
Sp2 Fa2 Consumption	0.95		Sp2 Fa2 Consumption	0.95	
Sp2 Fd3 Consumption	0.015		Sp2 Fd3 Consumption	0.015	
Sp3 O2 Consumption	75		Sp3 O2 Consumption	75	
Sp3 Fd1 Consumption	0.035		Sp3 Fd1 Consumption	0.035	
Sp3 Fa2 Consumption	0.015		Sp3 Fa2 Consumption	0.015	
Sp3 Fd3 Consumption	0.95		Sp3 Fd3 Consumption	0.95	
NPP	163	NPP	25731		
Evenness	0.62	Evenness	1.00		
Total Pairs	3	Total Pairs	305371		
Total Sp1 to Sp2 Pairs	1	Total Sp1 to Sp2 Pairs	110554		
Total Sp1 to Sp3 Pairs	1	Total Sp1 to Sp3 Pairs	39078		
Total Sp2 to Sp3 Pairs	1	Total Sp2 to Sp3 Pairs	38727		
d(Sp1-Sp2)	0.0738	d(Sp1-Sp2)	0.0738		
d(Sp1-Sp3)	0.0531	d(Sp1-Sp3)	0.0531		
d(Sp2-Sp3)	0.0329	d(Sp2-Sp3)	0.0329		
Dissimilarity	0.2937	Dissimilarity	0.3158		
Functional Diversity Index	0.1807	Functional Diversity Index	0.3151		
Base Scenario		Max Scenario			

Figure 7. Simulation results showing the baseline and maximum ecosystem configurations used to establish the range of the FDI.

Regression Analysis

Multiple regression analysis was used in order to better determine which ecosystem variables had the strongest prediction of variation in FDI within the simulated ecosystems. The model helped explain the variation, with an R-squared value of 84.13%, and an adjusted R-squared value 81.45%, which shows it's a strong predictor in the simulated model (internal validation in the software). (**Figure 8**). Most statistically significant predictors observed, which were p values less than 0.001, were most seen in the dietary composition variables such as Species2_Food2, showing a strong positive relationship with FDI at a coefficient value of +0.217, and Species3_Food2 showing a strong negative relationship with FDI at a coefficient value of -0.195 (**Figure 8**). These observed results show that the variation happening in food resource allocation to different species is a major contributor to variation in FDI within the simulated model. However, many variables are also able to show close significant trends, which include the Species2_Food1 at a p value around 0.068 and Species2_O2 consumption at a p value around 0.087 (**Figure 8**). However, the species count and oxygen consumption-related variables were not statistically significant predictors, as seen with the p-value being greater than 0.10. Additionally, these results furthermore show that functional diversity in the model is influenced more strongly by how species use resources than by differences in abundance. The variation inflation factor, which is also known as VIF, is extremely high in some abundance-related

variables. This indicates the strong multicollinearity among correlating species metrics. Therefore, it was concluded that the individual coefficients for these variables may be less reliable, but the holistic model still remains strong.

Figure 8. Regression model summary with coefficient analysis for identifying the most

Model Summary					
	S	R-sq	R-sq(adj)	R-sq(pred)	
	0.0053702	84.13%	81.45%	47.84%	
Coefficients					
Term	Coef	SE Coef	T-Value	P-Value	VIF
Constant	0.0421	0.0538	0.78	0.436	
Species1	0.000136	0.000081	1.68	0.097	200.47
Species2	0.000148	0.000213	0.70	0.489	1443.31
Species3	-0.000069	0.000360	-0.19	0.848	4016.06
gpp	-0.000001	0.000002	-0.31	0.761	12367.87
Species1_O2	-0.000414	0.000574	-0.72	0.473	1.47
Species1_Food1	-0.0239	0.0224	-1.07	0.289	3.92
Species1_Food2	0.0662	0.0614	1.08	0.284	4.02
Runs	0.00045	0.00107	0.42	0.676	1.49
Species2_O2	0.000452	0.000261	1.73	0.087	2.22
Species2_Food1	0.0998	0.0539	1.85	0.068	4.36
Species2_Food2	0.2173	0.0297	7.31	0.000	4.51
Species3_O2	-0.000459	0.000477	-0.96	0.339	1.43
Species3_Food1	0.0279	0.0205	1.36	0.177	1.44
Species3_Food2	-0.1948	0.0380	-5.12	0.000	1.37

statistically significant factors in predicting FDI variation.

Cart Threshold Analysis

In order to analyze the nonlinear ecosystem effects, a Classification and Regression Tree (CART) analysis occurred. The analysis showed that the Functional Diversity Index's variation was mainly affected by the two food composition variables of Species1_Food1 and Species2_Food2 (**Figure 9**). The CART analysis shows the clear threshold tipping points that can affect the functional diversity in the model. For example, if the Species1_Food1 variable is over 0.74, the FDI value experiences a drastic drop, which shows the functional degradation. However, in contrast, when the Species2_Food2 goes over 0.60, the FDI increases, which helps produce the highest FDI values as seen in Figure 8. Due to these thresholds, there are three distinct functional regimes: the baseline regime, which is a moderate FDI value with resources being used equally, a high performance regime, which is an increased FDI that is mainly influenced by the Species2_Food2 niches, and also a degradation regime, which is reduced FDI because of the overpresence of Species1_Food1. These findings illustrate that dietary composition is the main factor associated with variation in FDI within the simulated model. Because the variables examined in these analyzes also contribute to the formulation of the FDI, these results characterize the internal behavior of the index across randomized ecological configurations rather than providing independent ecological validation.

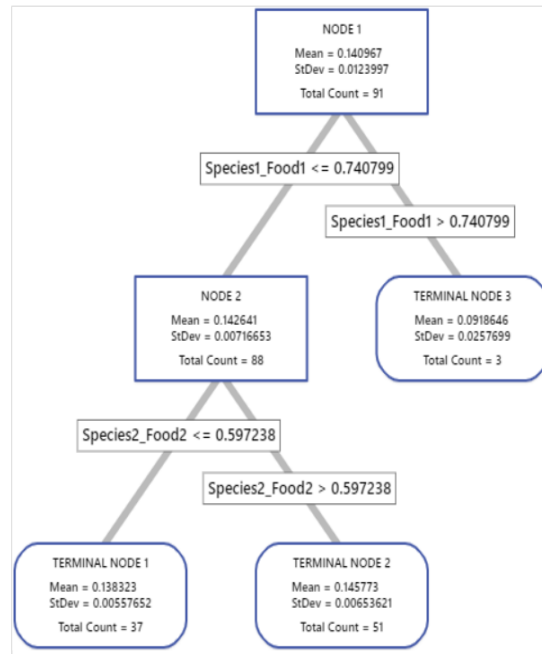


Figure 9. Trinodal CART tree showing non-linear thresholds for certain important variables. The values determine points where the ecosystem could perpetually destabilize.

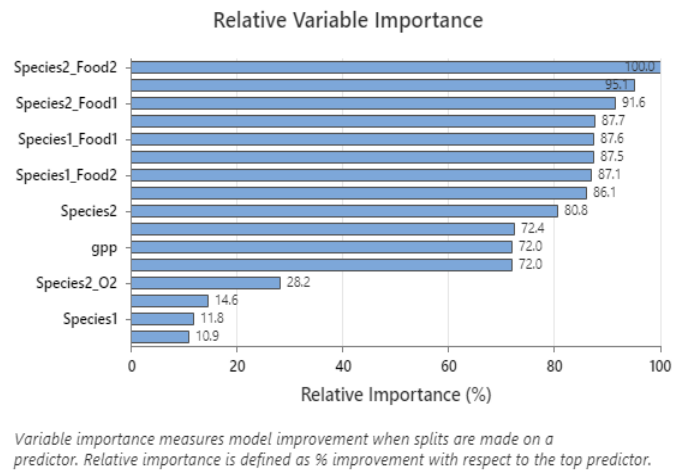


Figure 10. The graph covers the relative importance of simulative variables with respect to the top predictor of the functional diversity index.

K-Means Clustering Analysis

When representing complex ecological settings over multiple variable types, ensuring that the modeling systems are accurate and maximize representational utility is important. In terms of K-means clustering, it is crucial that the appropriate number of clusters is chosen to represent the data because issues like overfitting and underfitting can discard the data's applicability. Using a widely established method called the elbow method, which represents a point of diminishing returns, the optimal number of clusters identified was 3. When increasing the number of dimensions to practically large sets, the optimal cluster amount remained at 3, illustrating strength at that range for representation. In addition, the centroid aspect of this method was rudimentary for the intended type of centralized fluidity.

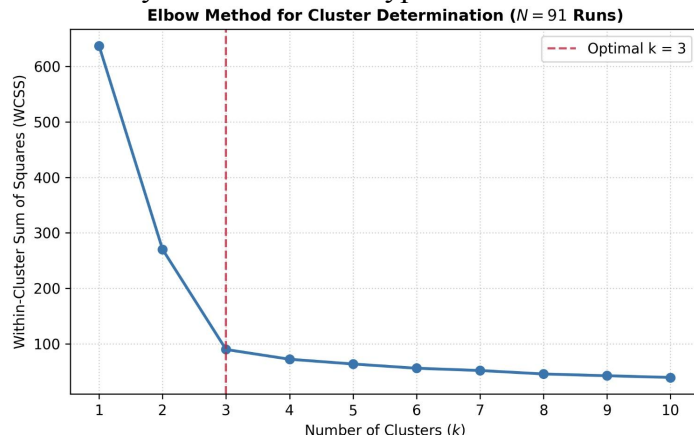


Figure 11. WCSS vs Number of Clusters (k). The elbow point indicates k=3 as the optimal amount.

Discussion

The results show that the FDI model is able to capture differences in modeled functional structure under the simulated conditions, which is based mainly on the consumption patterns of resources and O2-based productivity variables. Throughout the trials and simulation, it was seen that the index consistently was affected by the diet niche composition, which highlights how functional diversity seen in an ecosystem is based more on the ways organisms use resources rather than their sole species abundance.

Threshold-Driven Ecosystem Behavior

The CART analysis helped show how the ecosystem's responses were mainly following the nonlinear thresholds, but not the anticipated gradual changes. Also, when Species1_Food1 went over the value 0.74, the system behaved like a degradation regime in which it had a lower functional diversity in contrast to the high-performance regime achieved once Species2_Food2 was above the 0.60 value, subsequently giving the highest FDI value.

These results help support that the simulated ecosystem configurations can exhibit threshold transitions and distinct regimes within the FDI framework, where ecosystems can face volatility once those limits are reached. Traditional diversity indices, like the Shannon Diversity, continuously measure the changes within a population's diversity and therefore may miss the nonlinear transitions outlined in our model. Therefore, the given FDI identifies numeric thresholds within the simulated framework that may warrant further investigation using empirical ecological data.

Our FDI and other Biodiversity Metrics

Current biodiversity indices focus on taxonomic richness measures and evenness, omitting the functional roles that a species plays in its habitat. For example, Rao's Quadratic Entropy does have trait differentiation considered by them, oftentimes treating it as stable or stationary characteristics rather than as a variable for resource usage.

Our FDI model combines the O₂ consumption from GPP (gross primary productivity) with the dietary niche overlap, sourced from Pianka's Index. This new model helps connect the functional diversity of an ecosystem to the energy flow, which helps provide a holistic and also a more thorough understanding of ecosystem function when paired with the multiresource consumption that may also be prevalent.

Model Limitations

Although a randomized simulation procedure with 91 scenarios is widely regarded as statistically reliable, there are many limitations on what these results mean and can safely suggest based on the project prototype. Firstly, the method is purely an online simulation procedure over manually developed ecosystem ranges and attributes; thus, the findings here, although they may hint at broader truths, must be verified before real-life generalization. Concurring with that same idea, the model does not explicitly target more complex ecosystem functions like predation, reproduction, and migration, which hinders its usage in more surface-level contexts – contexts that don't have permissible access to oxygen and resource consumption rates. This is also seen in the fact that defining functionality by energy may narrow more complex ecological functions, like certain species being keystone or generalist populations. This means that further analysis is required to prove associations between the energy-weighted definitions and complex phenomena. In regards to limitations caused by the simulation assumptions, the most relevant is the fact that GPP and oxygen consumption are both simplified for traceability within the setting. That is, there is an assumed relationship between productivity and metabolism, which may be more complicated in real life.

Ecological Implications

The results of this study also suggest, within the simulated framework, FDI is more sensitive to how species use resources and their functional roles within the environment rather than to species abundance. Many traditional biodiversity indices, like the Shannon Diversity Index mentioned before, measure species richness and evenness, but may indicate that an ecosystem is diverse from the outside, but may not be functioning properly.

Being able to detect early on threshold shifts functionally, as seen in the CART analysis, scientists and ecologists could detect warning signs of an ecosystem about to collapse or degrade. In real life, these sudden changes in the functionalities of populations and species can help indicate the declining state of an ecosystem before species become endangered. With the current discussion of climate change, food security, and ecosystem stability, being able to detect the clear tipping point can help remove any ambiguity about when an ecosystem is about to fail.

Conclusion

Through multi-method statistical analysis and simulation testing, this study establishes a preliminary framework for evaluating functional diversity; under the tested and randomized conditions, the model variation was primarily driven by dietary resources rather than GPP dynamics. Additionally, the CART analysis illustrated that although variables may stay constant, their distribution among ecological populations is a key determinant in the functional diversity value. Moreover, the regression analysis clearly determined that the resource consumption variables had stronger associations with the functional diversity value, supporting the idea that dietary consumption is the most relevant factor. Additionally, as shown by the 91 random simulative distribution including the guiding base/max scenarios, the combination of Rao's Quadratic Entropy and Pianka's index jointly covered interspecies and intraspecies analysis, allowing for holistic ecosystem representation. This fact, in combination with the logical consistency and optimization model that demonstrates the functional diversity index's rudimentary behavior, supports the plausibility of using proposed frameworks to represent functional differences in increasingly complex ecosystems. Such models suggest ecological shifting between states rather than gradual change, which is a pattern that indices like the Shannon-diversity index overlook. Finally, the integration of K-means clustering used in marketing contexts into an ecological context allowed the model to be applied to various variable sets – sets that require multidimensional analysis and machine learning assistance. All in all, the empirical findings and algorithmic structures in this project may be vital in the contemporary era, classified by environmental degradation and biodiversity loss.

References

1. Botta-Dukát, Z. (2005). *Rao's quadratic entropy as a measure of functional diversity based on multiple traits*. *Journal of Vegetation Science*, 16, 533–540. <https://doi.org/10.1111/j.1654-1103.2005.tb02393.x>
2. 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. <https://doi.org/10.1890/03-9000>
3. Burbea, J., & Rao, C. R. (1982). *Entropy differential metric, distance and divergence measures in probability spaces: A unified approach*. *Journal of Multivariate Analysis*, 12(4), 575–596. [https://doi.org/10.1016/0047-259x\(82\)90065-3](https://doi.org/10.1016/0047-259x(82)90065-3)
4. Dakos, V., van Nes, E. H., D'Odorico, P., & Scheffer, M. (2012). *Robustness of variance and autocorrelation as indicators of critical slowing down*. *Ecology*, 93(2), 264–271. <https://doi.org/10.1890/11-0889.1>
5. Fattorini, S., Lombardo, P., Fiasca, B., Di Cioccio, A., Di Lorenzo, T., & Galassi, D. M. P. (2017). *Earthquake-related changes in species spatial niche overlaps in spring communities*. *Scientific Reports*, 7, 443. <https://doi.org/10.1038/s41598-017-00592-z>
6. Holling, C. S. (1973). *Resilience and stability of ecological systems*. *Annual Review of Ecology and Systematics*, 4(1), 1–23. <https://doi.org/10.1146/annurev.es.04.110173.000245>
7. Laliberté, E., & Legendre, P. (2010). *A distance-based framework for measuring functional diversity from multiple traits*. *Ecology*, 91(1), 299–305. <https://doi.org/10.1890/08-2244.1>
8. Lindeman, R. L. (1942). *The trophic-dynamic aspect of ecology*. *Ecology*, 23(4), 399–417. <https://doi.org/10.2307/1930126>
9. MacArthur, R., & Levins, R. (1967). *The limiting similarity, convergence, and divergence of coexisting species*. *The American Naturalist*, 101(921), 377–385. <https://doi.org/10.1086/282505>
10. Odum, E. P. (1969). *The strategy of ecosystem development*. *Science*, 164(3877), 262–270. <https://doi.org/10.1126/science.164.3877.262>
11. Pianka, E. R. (1974). *Niche overlap and diffuse competition*. *Proceedings of the National Academy of Sciences*, 71(5), 2141–2145. <https://doi.org/10.1073/pnas.71.5.2141>
12. Ricotta, C., & Szeidl, L. (2009). *Diversity partitioning of Rao's quadratic entropy*. *Theoretical Population Biology*, 76(4), 299–302. <https://doi.org/10.1016/j.tpb.2009.10.001>
13. Running, S. W., Nemani, R. R., Heinsch, F. A., Zhao, M., Reeves, M., & Hashimoto, H. (2004). *A continuous satellite-derived measure of global terrestrial primary production*. *BioScience*, 54(6), 547–560. [https://doi.org/10.1641/0006-3568\(2004\)054\[0547:ACSMOG\]2.0.CO;2](https://doi.org/10.1641/0006-3568(2004)054[0547:ACSMOG]2.0.CO;2)
14. Scheffer, M., Carpenter, S., Foley, J. A., Folke, C., & Walker, B. (2001). *Catastrophic shifts in ecosystems*. *Nature*, 413(6856), 591–596. <https://doi.org/10.1038/35098000>
15. Schoener, T. W. (1974). *Resource partitioning in ecological communities*. *Science*, 185(4145), 27–39. <https://doi.org/10.1126/science.185.4145.27>
16. Smouse, P. E., Banks, S. C., & Peakall, R. (2017). *Converting quadratic entropy to diversity: Both animals and alleles are diverse, but some are more diverse than others*. *PLOS ONE*, 12(10), e0185499. <https://doi.org/10.1371/journal.pone.0185499>