

MEASURING SPECIES DIVERSITY IN A SUBTROPICAL FOREST ACROSS A TREE SIZE GRADIENT: A COMPARISON OF DIVERSITY INDICES

XIANDONG KE, ZHIYAO SU*, YANQIU HU, YI ZHOU AND MINGFENG XU

College of Forestry and Landscape Architecture, South China Agricultural University, 510642 Guangzhou, China

*Corresponding author's e-mail: zysu@scau.edu.cn; Tel.: +86-20-85280263

Abstract

Shannon-Wiener index and Simpson's diversity index together with other metrics, e.g., richness, number of stems per species or species-specific density (N: S ratio), and kurtosis, were applied to characterize the woody plant diversity patterns of a subtropical broadleaved forest in south China. The aims of our study were to compare the efficacy and sensitivity to community diversity measures between Shannon-Wiener index and Simpson's diversity index. Tree census data from a 5-ha sample plot was partitioned into 3 datasets by diameter class to represent 3 distinct woody plant communities for the characterization of diversity across communities. The 5-ha sample plot of the forest had a total abundance of 23,301 tree stems ≥ 1 cm DBH and a richness of 139 species. The majority of tree stems were seedlings (41.1%) and saplings (38.8%), whereas canopy trees only accounted for 20.1% of the total tree stems. Both Shannon-Wiener index and Simpson's diversity index decreased significantly in response to a decrease in the N: S ratio across the datasets, but Shannon-Wiener index was more sensitive to plot-based richness changes and had a higher efficacy in predicting changes in species richness. Our findings are contrary to the general belief that Shannon-Wiener index is an insensitive measure of the character of the N: S relationship and have demonstrated that it remains a good measure for species diversity in plant community studies for its sensitivity and efficacy. We also suggest that the kurtosis statistic can be used as a new diversity measure due to its sensitivity to diversity change.

Key words: Shannon-Wiener index, Simpson's diversity index, Kurtosis, Size class, Subtropical forest.

Introduction

Species diversity is considered to be the key factor of ecosystem stability, complexity (Murdoch, 1975; Martinez *et al.*, 2006; Doebeli & Ispolatov, 2010), and ecosystem productivity (Hector, 2011; Omar *et al.*, 2016), due to its role as a crucial component of global diversity, and it is thus vital to the assessment of the ecosystem function (Whittaker, 1972; Bengtsson, 1998). Over the years, apart from using species richness and abundance to describe the structure and diversity of communities, a plethora of diversity indices have been proposed as measures of biological diversity (Magurran, 2004). Shannon-Wiener index and Simpson's diversity index are the two most popular diversity indices of all the proposed indices. Shannon-Wiener index, also known as Shannon Entropy, is a measure of information content and unpredictability in ecological studies (Margalef, 1958; Shannon *et al.*, 1949), whereas Simpson's diversity index describes the probability that a second individual drawn from a population should be of the same species as the first (Simpson, 1949). Both indices have been widely applied in fields such as ecological research, nature conservation, and natural resources assessment.

Although continued popularity has been enjoyed by Shannon-Wiener index, this ever-popular index was suggested to be not as good as its popularity would implies (Southwood & Henderson, 2000). It was believed to be a biased diversity measure dominated by abundant species and insensitive to the relationship or quantitative changes between number of stems and species richness, or N: S ratio, thus not being a good index for measuring species diversity (Yue, 1999; Southwood & Henderson, 2000). In contrast, Simpson's

diversity index was favored by some researchers for its simple equation and was regarded as sensitive to community changes (May, 1975; Simpson, 1949). Our study aims to compare the efficacy and sensitivity to community diversity measures between Shannon-Wiener index and Simpson's diversity index. For this purpose, we collected tree census data in a 5-ha subtropical evergreen broadleaved biodiversity monitoring plot. The census data were then partitioned into 3 datasets according to the diameter class to represent 3 woody plant communities with different N: S ratios (number of stems to number of species ratio). We also calculated kurtosis, a common statistic usually computed with a diversity measure, and evaluated its use in diversity measurements.

Materials and Methods

Study sites: Sampling plots were set up at Kanghe Provincial Nature Reserve (23°44'37"~23°52'16" N, 115°04'27"~115°09'41" E), which is located at the eastern part of South China's Guangdong Province. The Nature Reserve covers an area of 6,484.8 ha, running northeast to southwest. The region is semi-mountainous and semi-hilly with a complex and rugged topography and peaks in Baishigang with an elevation of 837.9 m. It is within the southern margin of the subtropical monsoon climate region and is characterized by an average annual temperature and precipitation of 21.1°C and 1 912 mm, respectively, and by a typhonic effect and intensive precipitation between April and September (Liang *et al.*, 2012). The area is mainly characterized by latosolic red soil, with mountain red soil, mountain yellow soil and meadow soil simultaneously. The forest community within the region houses a rich species

resource and is dominated by natural secondary evergreen broadleaved forest. The complex topography within the region offers various habitats for animal and plant survival and colonization. All the biotic resources are well protected, especially after the establishment of the nature reserve in October 2001.

Data collection: According to the investigation protocol of the tropical forest census plots (Condit, 1998), we established a 5-ha permanent sample plot in the Kanghe Provincial Nature Reserve. The plot is 250 m × 200 m and runs northwest to southeast, with northeast chosen as the major plot bearing. The plot was divided into 500 10 m × 10 m quadrats using a total station, and each corner of the 10 m × 10 m quadrat was marked with a PVC stake. The stake numbers were designated using 4 digits, starting from the left bottom at the major plot bearing. The first 2 digits, referring to the column number, are coded from west to east and begin from 00 to 25, and the last 2 digits, referring to the row number and are coded from south to north and begin from 00 to 20. Thus, the most south-western stake is 0000, and the most north-eastern stake is 2520. The plot code shares the same number with the stake of its south-western corner and thus ranges from 0000, indicating the most south-western plot, to 2419, indicating the most north-eastern plot facing the major plot bearing.

The forest inventory was processed based on the 10 m × 10 m quadrat. The DBH, species name, and corresponding information of all stems ≥ 1 cm DBH were measured and recorded. Electronic calipers were used for the measuring of trees ≤ 6 cm DBH, and diameter tapes for trees >6 cm DBH.

Datasets and statistical analysis: Plant census data in the 5-ha plot was classified into 5 diameter classes, namely seedlings, saplings, small trees, medium trees, and large trees, which were then designated to 3 different datasets. Dataset 1 includes all stems, dataset 2 includes all stems except seedlings, and dataset 3 excludes seedlings and saplings and includes all the canopy trees (Table 1).

The number of stems (*N*), species richness (*S*), number of stems per species, or species-specific density (*N*: *S* ratio, *N_s*), Shannon-Wiener index (*H'*), Simpson's

diversity index (*D'*), and kurtosis in each sample unit (10 m × 10 m quadrat) were calculated using the Row and Column Summary menu of PC-ORD (Version 6.0, MjM Software Design), a statistical software for ecological data analysis. In PC-ORD, Shannon-Wiener index and Simpson's diversity index are computed according to the following equations:

$$H' = -\sum_{i=1}^s P_i \ln P_i \tag{1}$$

$$D' = 1 - \sum_{i=1}^s P_i^2 \tag{2}$$

where *H'* = Shannon-Wiener index, *D'* = Simpson's diversity index, *s* = the species richness or number of species, and *P_i* = the proportional abundance of the *i*th species, in a sample unit. In PC-ORD, *D'* is calculated as a complement of Simpson's original index. It is the likelihood that two randomly chosen individuals will be different species.

Kurtosis measures the peakedness of distribution of species values (number of stems in a species) in a sample unit and is calculated in PC-ORD using the following equation:

$$Kurtosis = \left[\frac{n(n+1)}{(n-1)(n-2)(n-3)} \sum_{i=1}^n \left(\frac{x_i - \bar{x}}{s} \right)^4 \right] - \frac{3(n-1)^2}{(n-2)(n-3)} \tag{3}$$

where *s* = the sample standard deviation, *n* = the number of species, *x_i* = the number of stems of the *i*th species, \bar{x} = the number of stems per species or the mean species value, in a sample unit.

ANOVA (Analysis of Variance) was also performed together with regression analysis to compare the efficacy of the Shannon-Wiener index and Simpson's diversity index with respect to species richness, species-specific density, and kurtosis across the 3 datasets using Statistica (Version 8.0, StatSoft, Inc.).

Table 1. Species composition and diversity measured with different lower limits of size class.

Size class	DBH range (cm)	<i>N</i>	<i>S</i>	Dataset 1	Dataset 2	Dataset 3
Seedlings	1~2.4	9583	110	√		
Saplings	2.5~9.9	9028	111	√	√	
Small trees	10.0~19.9	3096	70	√	√	√
Medium trees	20.0~29.9	1198	35	√	√	√
Large trees	≥ 30.0	396	24	√	√	√
				<i>N</i> = 23,301	<i>N</i> = 13,718	<i>N</i> = 4,690
Size of dataset				<i>S</i> = 139	<i>S</i> = 123	<i>S</i> = 75
				<i>N_s</i> = 167.63	<i>N_s</i> = 111.53	<i>N_s</i> = 62.53

N = Number of stems; *S* = Number of species; *N_s* = Number of stems per species (*N*: *S*)

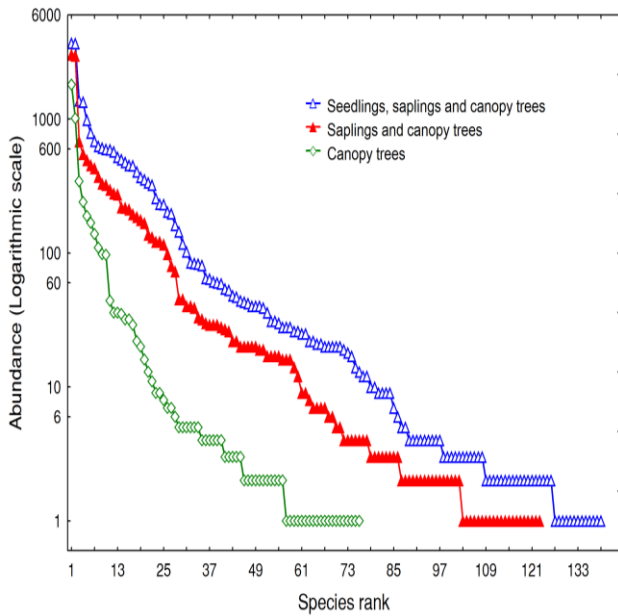


Fig. 1. Rank/abundance plot showing dominance patterns by size class.

Results and Discussion

Species composition and diversity: A total of 23,301 tree stems ≥ 1 cm DBH, representing a richness of 139 species, were recorded in the 5-ha sample plot. The majority of tree stems were seedlings (41.1%) and saplings (38.8%), and canopy trees only accounted for 20.1% of the total. Dataset 1 had the highest species-specific density and was mainly comprised of seedlings and saplings. Seedlings and saplings dominated 79.1% and 79.9%, respectively, of its species richness, and 41.1% and 38.8%, respectively, of its number of stems, whereas canopy trees only accounted for 20.1% of its number of stems. Saplings were also the major component of dataset 2 and occupied 90.2% and 65.8% of its species richness and number of stems, respectively. In contrast, dataset 3 had the lowest species-specific density, and small trees constituted 93.3% and 66% of its species richness and number of stems, respectively. Up to 16 plant species occurred only as seedlings and 48 plant species as seedlings or saplings, resulting in 64 plant species found only as seedlings or saplings, which accounted for 46% and 79.9% of the total species richness and number of stems, respectively (Table 1).

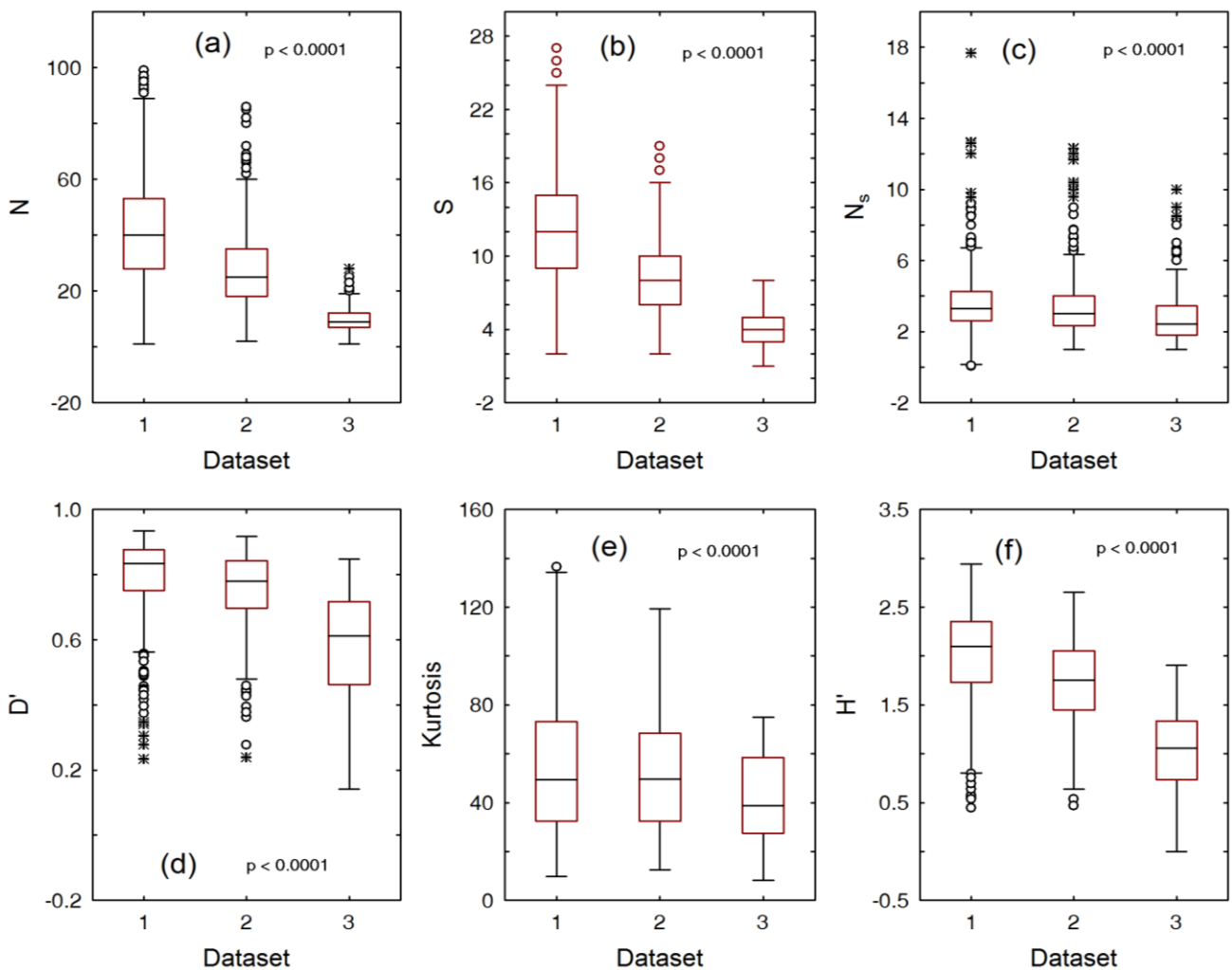


Fig. 2. Variation in diversity measures across datasets. The horizontal line in each box indicates the median, the box endpoints are the 25th and 75th percentile values, and the whiskers represent the non-outlier range, with the circles indicating the outliers and the asterisks the extreme values of a particular variable. The p -value was obtained using the Kruskal-Wallis test. Abbreviations: N = Number of stems; S = Number of species; N_s = Number of stems per species; D' = Simpson's diversity index; H' = Shannon-Wiener index.

The rank/abundance curves of the 3 datasets, which all followed a log series distribution (Fisher *et al.*, 1943), revealed that all 3 datasets hold more rare species and less abundant species (Fig. 1). The above results showed that the 3 datasets consisting of stems from different diameter classes were all dominated by diameter classes with a lower DBH and that all 3 datasets also had a significantly different N: S ratio.

The ANOVA results illustrated in Fig. 2 demonstrated that the extreme significance ($p < 0.0001$) existing in the number of stems, species richness, species-specific density, Shannon-Wiener index, Simpson's diversity index, and kurtosis could be caused by the different species compositions and various abundance distributions across the 3 datasets. The values of the above 6 indices ranked in the 3 datasets were shown as dataset 1 > dataset 2 > dataset 3. The inclusion of seedlings and saplings appeared to enable the community to have a higher species richness and species diversity, which may be because, in contrast to the community of classes with a larger

diameter, the seedling and sapling communities were primarily affected by the quality and quantity of the seeds. Other than resource competition, regeneration niche (Grubb, 1977) dominated the regeneration and colonization processes, thus eventually affecting their occurrences (Svenning *et al.*, 2008). All of these factors permitted the coexistence of more species and individuals and resulted in a higher species richness and number of stems in the seedling and sapling communities. As seedlings and saplings grew up into a larger diameter class, the effect of predators and of interindividual competition for resources was also increased accordingly. In addition, a negative density-dependent mechanism (Wills *et al.*, 1997; Wright, 2002) alternatively governed the larger diameter community, causing considerable death of stems to balance the relationship between the environment and community. The death of stems caused a decrease in the species and number of stems and ultimately led to a lower species richness and species diversity.

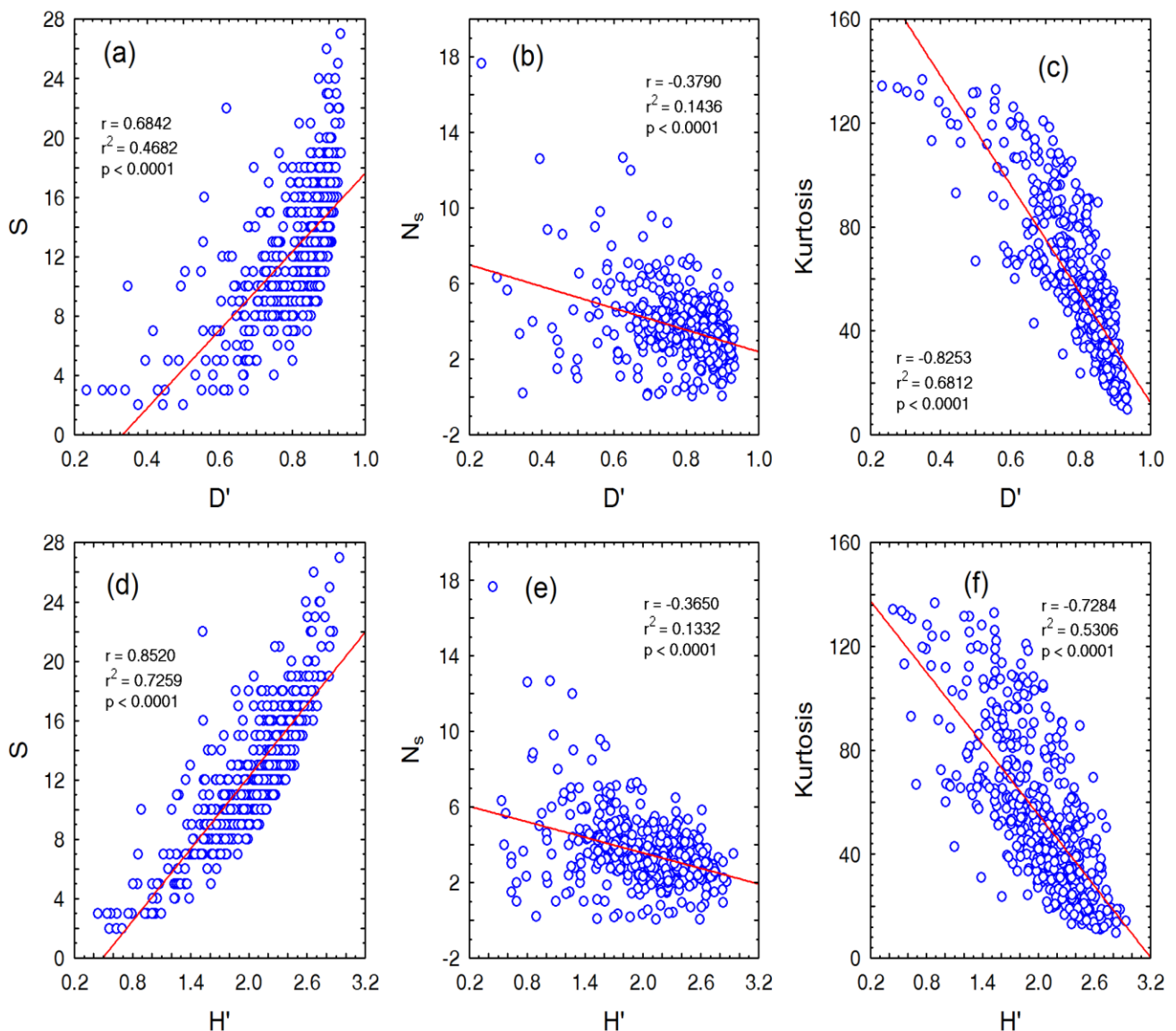


Fig. 3. Efficacy of Shannon-Wiener index and Simpson's diversity index in predicting diversity changes in plots for dataset 1. Abbreviations: S = Number of species; N_s = Number of stems per species; D' = Simpson's diversity index; H' = Shannon-Wiener index.

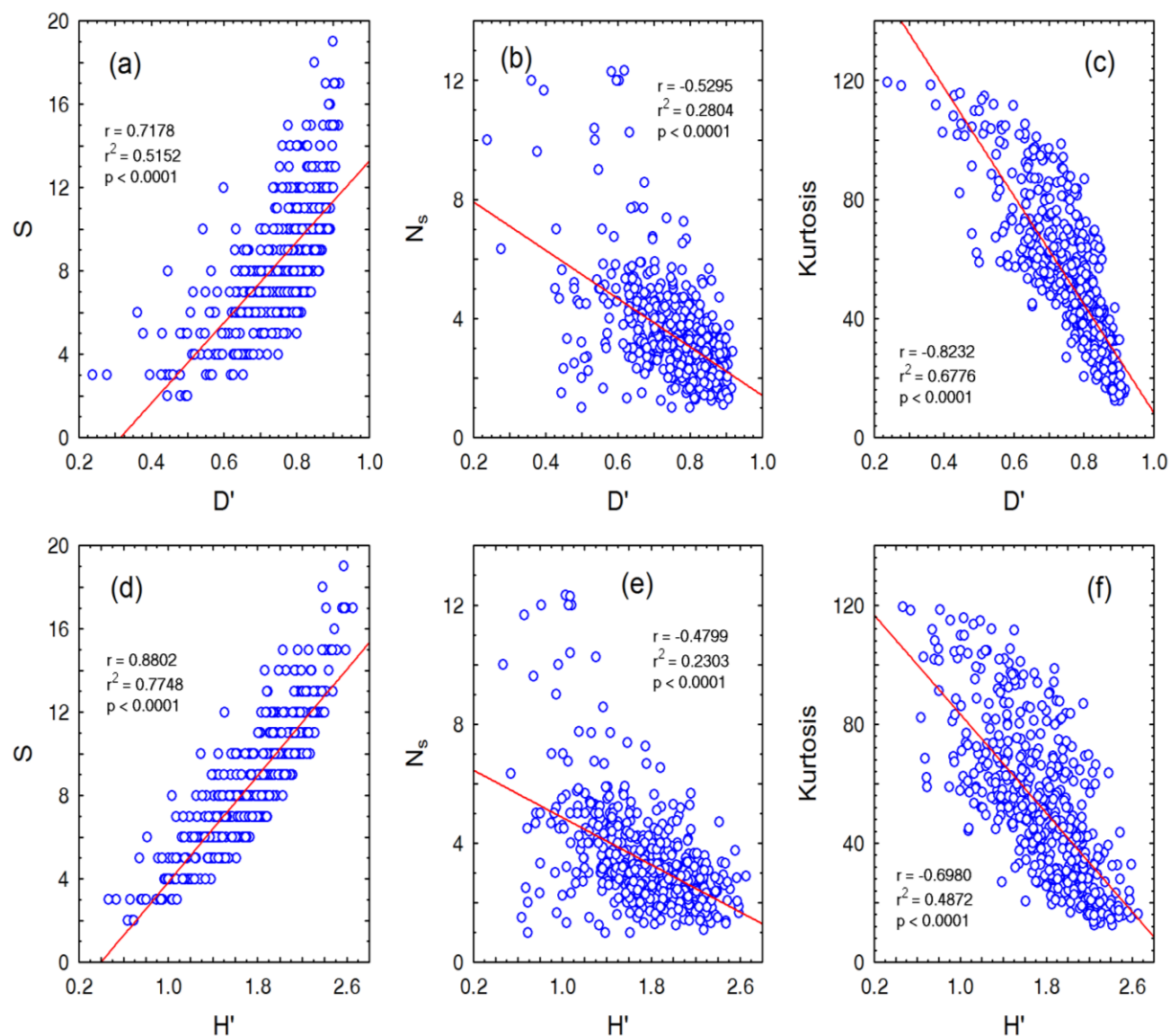


Fig. 4. Efficacy of Shannon-Wiener index and Simpson's diversity index in predicting diversity changes in plots for dataset 2. Abbreviations: S = Number of species; N_s = Number of stems per species; D' = Simpson's diversity index; H' = Shannon-Wiener index.

Sensitivity of Shannon-Wiener index and the Simpson's diversity index: Both Shannon-Wiener index and the Simpson's diversity index showed a significantly positive correlation ($r > 0.68$, $p < 0.0001$) with species richness in 3 datasets. The determination coefficient r^2 in the 3 datasets was ranked as follows: dataset 1 (Fig. 3, a & d) < dataset 2 (Fig. 4, a & d) < dataset 3 (Fig. 5, a & d). In addition, the Shannon-Wiener index was more sensitive than the Simpson's diversity index to species richness changes and was more capable of reflecting the species composition.

Similarly, a significantly negative correlation ($r < -0.69$, $p < 0.0001$) was detected between both the Shannon-Wiener index and the Simpson's diversity index and kurtosis in the 3 datasets. The ranking of the determination coefficient r^2 in the 3 datasets was slightly different from the former situation and was revealed to be dataset 2 (Fig. 4, c & f) < dataset 1 (Fig. 3, c & f) < dataset 3 (Fig. 5, c & f). The Simpson's diversity index was more sensitive than Shannon-Wiener index to

kurtosis changes and was capable of reflecting the species distribution.

Although neither the Shannon-Wiener index nor the Simpson's diversity index could provide an robust explanation of the variation of species-specific density ($r^2 < 0.29$), which could suggest that both diversity indices were not sensitive, they were found to be significantly negatively correlated ($r < 0.36$, $p < 0.0001$) with the species-specific density (Fig. 3, b & e; Fig. 4, b & e; Fig. 5, b & e). The Shannon-Wiener index and the Simpson's diversity index both efficiently reflected the changes in species richness, kurtosis, and species-specific density. Although a few other diversity measures offered by some researchers were believed to be more thorough and have been suggested as substitutes for the Shannon-Wiener index, we insisted that a good diversity measure should not only be sensitive but also simple to use, two factors that therefore made the Shannon-Wiener index a superior diversity measure.

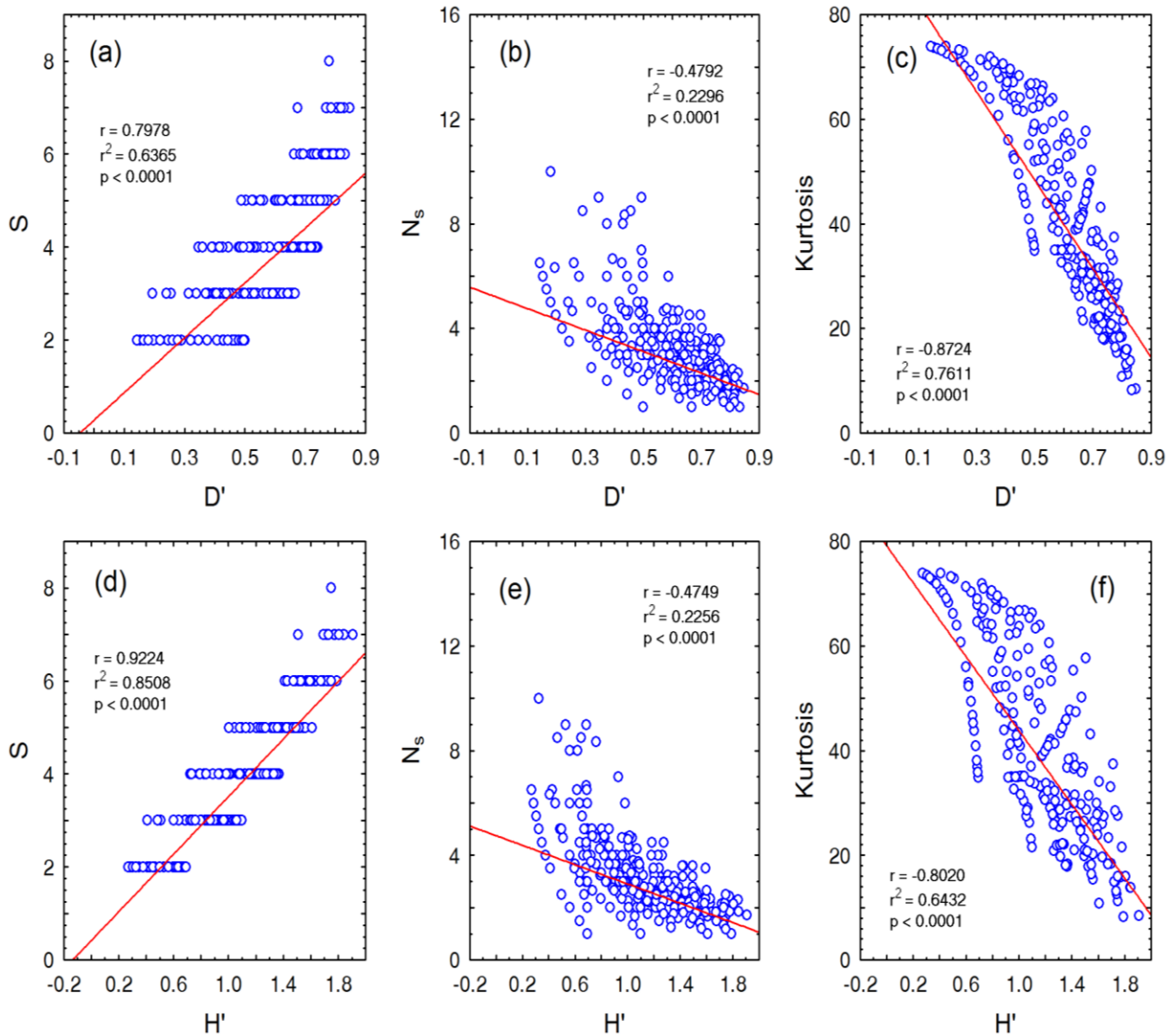


Fig. 5. Efficacy of Shannon-Wiener index and Simpson's diversity index in predicting diversity changes in plots for dataset 3. Abbreviations: S = Number of species; N_s = Number of stems per species; D' = Simpson's diversity index; H' = Shannon-Wiener index.

Conclusions

It is from the study concluded that both Shannon-Wiener index and Simpson's diversity index are sensitive to the changes in species richness and kurtosis. However, Shannon-Wiener index is more sensitive in predicting species composition changes. The two diversity indices are significantly and negatively correlated with the number of stems per species, or species-specific density (N: S ratio).

N: S relationship also reflects the relative percentage of the number of rare species. Given a certain amount of individuals, the higher the N: S value, the fewer the rare species are included. Our findings are contrary to the assertion that Shannon-Wiener index is an insensitive measure of the character of the N: S relationship, indicating that the Shannon-Wiener index might as well be sensitive to the rare species. The results have demonstrated that the Shannon-Wiener index remains a good measure for species diversity in plant community studies, at least as good as Simpson's diversity index with respect to sensitivity and efficacy. The kurtosis statistic,

which involves the information of abundance distribution pattern, was sensitive to diversity changes. A lower kurtosis value reflects a more average plot-based abundance across species, thus representing higher species diversity. Therefore, the kurtosis statistic can be used as a new diversity measure, especially in stand-level plant community analysis.

Acknowledgements

The authors would like to thank Dr. Zhenkui Li and Mr. Wenbin Li for their help in field plant identification. This research is supported by the Science and Technology Plan Programs of Guangdong Province (Grant No. 2013B020305008) and a special grant from Guangdong Forestry Administration.

References

- Bengtsson, J. 1998. Which species? What kind of diversity? Which ecosystem function? Some problems in studies of relations between biodiversity and ecosystem function. *Appl. Soil Ecol.*, 10(3): 191-199.

- Condit, R. 1998. *Tropical forest census plots: methods and results from Barro Colorado Island, Panama and a comparison with other plots*. Springer, Texas, USA.
- Doebeli, M. and I. Isolatov. 2010. Complexity and diversity. *Science*, 328(5977): 494-497.
- Fisher, R.A., A.S. Corbet and C.B. Williams. 1943. The relation between the number of species and the number of individuals in a random sample of an animal population. *J. Ani. Ecol.*, 12: 42-58.
- Grubb, P. J. 1977. The maintenance of species-richness in plant communities: The importance of the regeneration niche. *Biol. Rev.*, 52(1): 107-145.
- Hector, A. 2011. Diversity favours productivity. *Nature*, 472(7341): 45-46.
- Liang, B.H., Q.R. Pan and C.M. Wei. 2012. An Investigation of wild fruit plants resources in Dongyuan Kanghe Provincial Nature Reserve, Guangdong Province. *Subtropical Plant Science*, 41(3): 56-60.
- Magurran, A.E. 2004. *Measuring biological diversity*. Blackwell Science, Oxford, UK.
- Margalef, R. 1958. Information theory in ecology. *General Systems: Yearbook of the International Society for the Systems Sciences*, 3: 36-71.
- Martinez, N.D., R.J. Williams and J.A. Dunne. 2006. Diversity, complexity, and persistence in large model ecosystems. In: (Eds.): Pascual, M. and J.A. Dunne. *Ecological Networks: Linking Structure to Dynamics in Food Webs*. Oxford University Press, New York, pp. 163-185.
- May, R.M. 1975. Patterns of species abundance and diversity. In: (Eds.): Cody, M.L. and J.M. Diamond. *Ecology and Evolution of Communities*. Harvard University Press, Cambridge, USA, pp. 81-120.
- Murdoch, W. W. 1975. Diversity, complexity, stability and pest control. *J. Appl. Ecol.*, 12(3): 795-807.
- Omar, M.Y., A. Maroyi and J.J. Van Tol. 2016. Floral diversity, composition and distribution in a montane wetland in Hogsback, the eastern Cape Province, South Africa. *Pak. J. Bot.*, 48(5): 1861-1870.
- Shannon, C.E., W. Weaver, R.E. Blahut and B. Hajek. 1949. *The mathematical theory of communication*. University of Illinois Press, Urbana, USA.
- Simpson, E.H. 1949. Measurement of diversity. *Nature*, 163: 688.
- Southwood, T.R.E. and P.A. Henderson. 2000. *Ecological methods*. (3rd Ed) Blackwell Science, Oxford, UK.
- Svenning, J.C., T. Fabbro and S.J. Wright. 2008. Seedling interactions in a tropical forest in Panama. *Oecologia*, 155(1): 143-150.
- Whittaker, R.H. 1972. Evolution and measurement of species diversity. *Taxon*, 21: 213-251.
- Wills, C., R. Condit, R.B. Foster and S.P. Hubbell. 1997. Strong density- and diversity-related effects help to maintain tree species diversity in a neotropical forest. *Proceedings of the National Academy of Sciences*, 94(4): 1252-1257.
- Wright, J.S. 2002. Plant diversity in tropical forests: A review of mechanisms of species coexistence. *Oecologia*, 130(1): 1-14.
- Yue, T.X. 1999. Studies on models for Biodiversity. *J. Nat. Resources*, 14(4): 377-380.

(Received for publication 12 June 2016)