

## **Comparative analysis on ecosystem diversity indices using SAS computer programming**

(Penggunaan pengaturcaraan komputer SAS dalam analisa perbandingan indeks kepelbagaian ekosistem)

M. M. Rashid\*, S. T. S. Hassan\*, I. Azhar\*\* and A. R. Ismail\*

Key words: diversity indices, biodiversity, SAS computer program, paddy arthropods, comparative analysis.

### **Abstrak**

Pengaturcaraan SAS ('Statistical Analysis System') yang lengkap dikemukakan untuk digunakan dalam pengiraan indeks kepelbagaian biologi terhadap beberapa atropod sawah padi. Analisis varians digunakan dalam meneliti kepentingan parameter (pembolehubah) utama terhadap indeks kepelbagaian, diikuti dengan ujian perbandingan bererti antara parameter tersebut bagi setiap indeks. Kesan yang bererti didapati antara corak penanaman (secara tanam ubah dan tabur terus) berdasarkan indeks kekayaan, kepelbagaian dan keserataan. Walau bagaimanapun, corak penanaman tidak mempengaruhi indeks siang-malam kecuali dalam indeks kekayaan bagi spesies, N0 dan R1. Nilai indeks kepelbagaian untuk keseluruhan ekosistem terletak di antara nilai-nilai indeks bagi kedua-dua corak penanaman tersebut. Corak penanaman secara tanam ubah menunjukkan kekayaan dan kepelbagaian yang tinggi. Data yang banyak boleh digunakan secara terus dengan penggunaan pengaturcaraan ini. Pengguna bebas meneliti kesan pembolehubah dengan memilih atau mengeluarkan parameter yang berkaitan. Indeks untuk parameter yang berbeza boleh dikira dan dianalisa secara serentak dan berterusan.

### **Abstract**

A comprehensive SAS (Statistical Analysis System) programming is proposed to be used in the calculation of biodiversity indices on paddy arthropods, as an example. An ANOVA was used in testing the significance effect of major parameters on diversity indices, followed by significant comparison testing between those parameters for each index. There are significant effects between type of planting (transplanted and direct-seeded) on richness, diversity and evenness indices. However, type of planting did not influence day and night diversity indices, except in species richness, N0 and R1. Diversity indices for the entire ecosystem were intermediate in values between those of the two types of planting. The transplanted field indicates higher richness and diversity. Novel features presented in this programme are; a large data set can be accessed directly with this program, and the user is free to explore variable effects through selection or elimination of the parameter(s) concerned. Hence, the various diversity indices can be calculated and analysed for different parameters simultaneously and continuously.

---

\*Department of Biology, Universiti Putra Malaysia, 43400 UPM, Serdang, Selangor, Malaysia

\*\*MARDI Research Station, Hilir Perak, P. O. Box 25, 36307 Sungai Sumun, Perak, Malaysia

Authors' full names: Mansor Mohd Rashid, Syed Tajuddin Syed Hassan, Azhar Ismail and Abdul Rahim Ismail

©Malaysian Agricultural Research and Development Institute 1999

## Introduction

Ecosystem diversity is an important descriptor of ecosystem dynamics through interaction or restriction of components or species, by flows of energy, materials, and information (Jizhong et al. 1991). The species richness, species abundance and the interacting patterns among the components lead to the diversity or biodiversity of ecosystem. In ecological research, ecosystem diversity is a major area because of its relations to the stability and productivity of an ecosystem (Jizhong et al. 1991). Recently, index of diversity was used in comparative analysis of the ecosystem, e.g., analysis of land with regard to vegetation physiognomy (Gomez-Sal et al. 1993) and seasonal dynamics of birds (Farina 1993) due to the landscape transformation, using Shannon index. Moreover, several diversity indices were used on paddy arthropod populations in comparing between sites and planting patterns (Hassan and Rashid 1997).

Ecologists have proposed numerous indices of diversity to measure species richness, abundance and evenness (Shannon and Weaver 1949; Simpson 1949; Pielou 1969, 1975; Odum et al. 1960; Hill 1973; Peet 1974; DeJong 1975; Margalef 1968; MacIntosh 1967; Hurlbert 1971). Ludwig and Reynolds (1988) compiled some of these diversity indices together with a GWBASIC programme for their computation. Their programme is interactive, and requires the user to key-in the number of species observed and the number of individuals of the particular species. Unfortunately, their methodology is suitable only for a small sample size, and becomes impractical for large samples, with many sampling occasions or parameters. In the GWBASIC program of Ludwig and Reynolds (1988), the user needs to organize the data properly and predetermines the number of individuals and the number of species observed. Moreover, this programme was limited to handle only one parameter at a time. Hence, it is necessary to run programme repetitively and separately, in

order to analyse many parameters and different levels of parameters respectively.

Consequently, this paper introduced a SAS (Statistical Analysis System) programme, which can compute diversity indices for large samples simultaneously, many parameters and different levels of parameters. These indices were then used in comparative analysis of a wet paddy ecosystem diversity. Another novelty of this programme is that the raw data i.e. the number of individuals of a particular species, can be input directly into the SAS programme structure. The optimal way of organizing the data before being used in the programme is also indicated. The validity of this SAS programme was further validated by running it using fish-catch data of Livingston (1976). Values of diversity indices generated were then compared with those calculated by Ludwig and Reynolds (1988) GWBASIC programme.

## Materials and methods

### *Data source*

Data were collected from two rice plots with different planting techniques (direct seeding and transplanting) at Sawah Sempadan, Tanjung Karang, Selangor (3° 20' N, 101° 12' E) in 1992. Each plot (67 m x 61 m) was established with MR84 variety. Direct seeding was done on 2 February 1992, and transplanting on 17 February 1992, 21 days after seeding in the nursery. At each plot, direct visual counts on 22 categories of arthropod were conducted weekly at 3 hourly intervals, during 24 h duration. Sampling commenced on 23 April through 3 June 1992 using one hill (for transplanted) and one naturally-formed clump (for direct seeding) as sampling unit, and 20 hills were examined at random. At each site, the field was sampled by walking through the field with a zig-zag or diagonally cross pattern to maximize coverage.

A fish catch data set for a north Florida estuary (Livingston 1976) was then used to validate the SAS programming algorithm. The diversity indices generated were then

compared to those reported by Ludwig and Reynolds (1988).

**Data organization**

A SAS® (SAS 1990) environment facilitates entering of data either directly or indirectly into its programming structure. The direct method caters for a small data set. For a large set of data comprising many parameters/variables, with more than two levels for each parameter/variable, the indirect method is recommended. Lotus® software (Lotus 1-2-3 1993) was used in this study to organize and manage the data, due to its user-friendly and easy to use features. Lotus data files were converted into print format file before transferring to SAS environment (Figure 1). As an example, our data comprises five parameters i.e. planting pattern (2 types), date of sampling (7 dates for each planting type), species code (41

species within the 22 categories), time of sampling (8 times within 24 h) and replicates (20 hills). Thus there are 91 840 data points. Table 1 shows a Lotus worksheet data for direct-seeded paddy, as an example. Other software that permits the user to input data is Microsoft Excel that allows the text format conversion. Note that the print format files of Lotus or text format files of Microsoft Excel can be accessed directly in SAS environment.

**Development of SAS programme algorithm and analysis**

In developing the SAS algorithm, indices of diversity comprising components of species richness, evenness and diversity indices (Ludwig and Reynolds 1988) were used (Table 2). The validity of this SAS programme (Appendix 1) was tested using fish catch data of Livingston (1976). Values of diversity indices generated were then compared with those calculated by Ludwig and Reynolds (1988) using their GWBASIC programme.

A SAS programming structure (Appendix 2) was also developed and used to compute the diversity indices based on differing dates of sampling, times of sampling, day-night samplings, types of planting pattern and the entire duration (all sampling occasions) on paddy arthropods. In comparing the statistical differences between planting pattern, a one-way ANOVA on each diversity index followed by mean comparisons were conducted using the PROC GLM of SAS (SAS 1990).

The structural programme consists of three sections A, B and C (Appendix 2). Data input is placed in the first section (A). The first row of each section must be initialized and provided with a project name. The name DIVER is used here to represent the raw data. The INFILE, EOF (End Of File) and RETURN statements allow processing of many data files. Here, there are four data files (filename1 through filename 4) with six variables i.e. type (type of planting), date (date of sampling), time

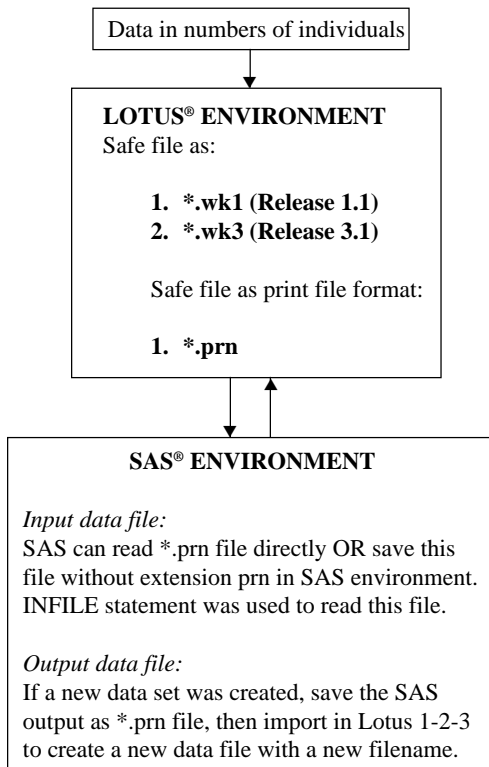


Figure 1. General algorithm in optimally organizing large data sets before running the SAS program

Table 1. An example of Lotus worksheet for data input that are saved as print file format

DS	1	1	1	1	0	DS	2	1	1	1	0	DS	3	1	1	1	0
DS	1	1	1	2	0	DS	2	1	1	2	3	DS	3	1	1	2	2
DS	1	1	1	3	0	DS	2	1	1	3	7	DS	3	1	1	3	0
DS	1	1	1	4	0	DS	2	1	1	4	0	DS	3	1	1	4	0
DS	1	1	1	5	3	DS	2	1	1	5	0	DS	3	1	1	5	3
DS	1	1	1	6	1	DS	2	1	1	6	0	DS	3	1	1	6	1
DS	1	1	1	7	0	DS	2	1	1	7	6	DS	3	1	1	7	0
DS	1	1	1	8	4	DS	2	1	1	8	1	DS	3	1	1	8	6
DS	1	1	1	9	0	DS	2	1	1	9	1	DS	3	1	1	9	0
DS	1	1	1	10	0	DS	2	1	1	10	2	DS	3	1	1	10	4
DS	1	1	1	11	3	DS	2	1	1	11	1	DS	3	1	1	11	4
DS	1	1	1	12	0	DS	2	1	1	12	0	DS	3	1	1	12	6
DS	1	1	1	13	0	DS	2	1	1	13	2	DS	3	1	1	13	0
DS	1	1	1	14	2	DS	2	1	1	14	6	DS	3	1	1	14	3
DS	1	1	1	15	0	DS	2	1	1	15	0	DS	3	1	1	15	2
DS	1	1	1	16	2	DS	2	1	1	16	0	DS	3	1	1	16	5
DS	1	1	1	17	0	DS	2	1	1	17	1	DS	3	1	1	17	0
DS	1	1	1	18	0	DS	2	1	1	18	2	DS	3	1	1	18	0
DS	1	1	1	19	0	DS	2	1	1	19	8	DS	3	1	1	19	6
DS	1	1	1	20	2	DS	2	1	1	20	0	DS	3	1	1	20	1
DS	1	2	1	1	0	DS	2	2	1	1	3	DS	3	2	1	1	0
DS	1	2	1	2	2	DS	2	2	1	2	7	DS	3	2	1	2	4
.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
DS	1	8	41	18	0	DS	2	8	41	18	2	DS	3	8	41	18	4
DS	1	8	41	19	0	DS	2	8	41	19	0	DS	3	8	41	19	8
DS	1	8	41	20	0	DS	2	8	41	20	0	DS	3	8	41	20	2

Note: Each column was repeated three times with the sequence: Type of planting (DS – direct-seeded); Date of sampling; Time of sampling (1–8); Species code (1–41); Hill number (replication) (1–20); Number of individuals observed. Each column consists of 6560 rows for each sampling occasion (date).

(time of sampling), *sppcode* (species code - numeric), *hillno* (hill number as replicate), and *num* (number of individuals observed). The *hillno* can be made unselected using DROP statement.

In the second section (B), the algorithm is to calculate the biodiversity indices, either by each variable or by a group of variables. If a group of variables was chosen, then biodiversity indices can be compared. A SET statement is used to recall the data DIVER in section A. The DROP statement is subsequently used to exclude the unimportant variable(s). The PROC NPAR1WAY and PROC GLM statements were used to assess the significance of biodiversity indices calculated at different major parameters (i.e. planting type, date

and time of sampling). The former procedure statement is used to execute the non-parametric test of the indices through  $\chi^2$  of Kruskal-Wallis test. Since each mean data value of the indices (*Table 1* and *Table 2*), represents 56 data values, the Central Limit Theorem operates, thus ensuring normality of the data values. Subsequently the parametric test of the PROC GLM statement is used to execute the analysis of variance and mean comparison using Ryan-Einot-Gabriel-Welsch Multiple Range Test (REGWQ) for data with equal sample sizes and TUKEY or TUKEY-WALLER for unequal sample size (SAS 1990). Processing of this section can be made interactive involving different variable(s) and varying combinations, continually.

Table 2. List of diversity indices studied by Ludwig and Reynolds (1988) on fish catches data of Livingston (1976)

Formula	Emphasis of index	References
$R1 = (S-1)/\ln(n)$	Richness	Margalef (1968)
$R2 = S/\sqrt{n}$	Richness	Menhinick (1964)
$N0 = S$	Richness	Hill (1973)
$N1 = e^{H'}$	Diversity	Hill (1973)
$N2 = 1/\lambda$	Diversity	Hill (1973), Simpson (1949)
$\lambda = \sum p_i^2$	Diversity	Simpson (1949)
$H' = -\sum (p_i \ln p_i)$	Diversity	Shannon and Weaver (1949)
$E1 = H'/\ln S = \ln N1/\ln N0$	Evenness	Pielou (1975)
$E2 = e^{H'}/S = N1/N2$	Evenness	Sheldon (1969)
$E3 = (e^{H'}-1)/(S-1) = (N1-1)/(N0-1)$	Evenness	Heip (1974)
$E4 = 1/\lambda / e^{H'} = N2/N1$	Evenness	Hill (1973)
$E5 = [(1/\lambda)-1]/(e^{H'}-1) = (N2-1)/(N1-1)$	Evenness	Alatalo (1981)

$S$  – the total number of species;  $n$  – the total number of individuals observed;  $p_i$  – the proportional abundance of the  $i$ th species,  $p_i = n_i/N$  where  $n_i$  is the number of individuals of the  $i$ th species and  $N$  is the total number of individuals for all  $S$  species in the population.

Table 3. SAS printed output on diversity indices of fish data of Livingston (1976) were calculated using SAS program in *Appendix 1*. The results shown were similar to those calculated and presented by Ludwig and Reynolds (1988)

STATISTICAL ECOLOGY						
9:30 Thursday, August 24, 1997						
Diversity indices for December collections						
ND	N0D	R1D	R2D	DLAMDA	HD	N1D
761	20	2.86376	0.72500	0.22772	1.91351	6.77680
N2D		E1D	E2D	E3D	E4D	E5D
4.39143		0.63874	0.33884	0.30404	0.64801	0.58708
Diversity indices for June collections						
NJ	N0J	R1J	R2J	JLAMDA	HJ	N1J
1562	15	1.90380	0.37953	0.25616	1.54451	4.68569
N2J		E1J	E2J	E3J	E4J	E5J
3.90377		0.57034	0.31238	0.26326	0.83313	0.78785

Finally, the third section (C) is used to calculate the biodiversity indices for the entire system. This section is located at the end of a structured programme. As in section B, the SET statement is used to access the raw data. Variables can be unselected using DROP statement except *sppcode* and *num*. A new data set was created in this section to generate a correct calculation of the biodiversity indices.

### Results and discussion

The similarities of the biodiversity indices values generated from our SAS programming (*Table 3*) and those by Ludwig and Reynolds (1988) using Livingston's (1976) fish data, vindicates our SAS programme algorithms.

In the preliminary analysis, only 5 sampling dates of direct-seeded were used and the resulting biodiversity indices were compared to those of the 7 sampling dates of transplanted rice. The species richness ( $R1$  and  $R2$ ) differed between these two

sources of data when dates were chosen as a source of variability. A non-parametric statistics comparison using Kolmogorov-Smirnov Two-Sample Test (Statgraphics 1988) showed that they (planting types) were significantly different for both indices, with the maximum absolute deviation (DN) between their cumulative distribution functions (cdf) at 0.75 ( $p < 0.05$ ) for R1 (Figure 2a) and 0.88 ( $p < 0.01$ ) for R2 (Figure 2b). Obviously, these two indices generally depend on the number of individuals (N), and species number (N0), which increase with increasing sample size (Ludwig and Reynolds 1988). This indicates that in a comparative analysis, the unequal sample size of two different sites can lead to biased conclusions and interpretations. Therefore, additional two sampling dates of direct-seeded data were generated by simulation based on resampling technique to ensure orthogonality of variances. In the resampling method, random numbers were generated using MS-DOS QBASIC programming and were used to select samples randomly with replacement from the actual direct-seeded data sets.

In assessing the significance of diversity indices using ANOVA, with type of planting as the main effect, all diversity indices were significantly influenced by the main effect except for E4 and E5 of evenness and in Menhinick's species richness (R2) of date effect (Table 4). Similar results were obtained when using a non-parametric test of  $\chi^2$  statistics of Kruskal-Wallis test. It is worth noting that the number of individuals (N) were not significantly affected but the species number (N0) were significantly affected by the type of planting ( $p < 0.01$ ) (Table 4). Therefore, the contradicting result shown in species richness (R2) of date effect as compared with time effect was due to the sensitivity of this index to the species number (N0) (Ludwig and Reynolds 1988). In comparing direct-seeded and transplanted effects, type of planting has no significant effect on species diversity and evenness based on day

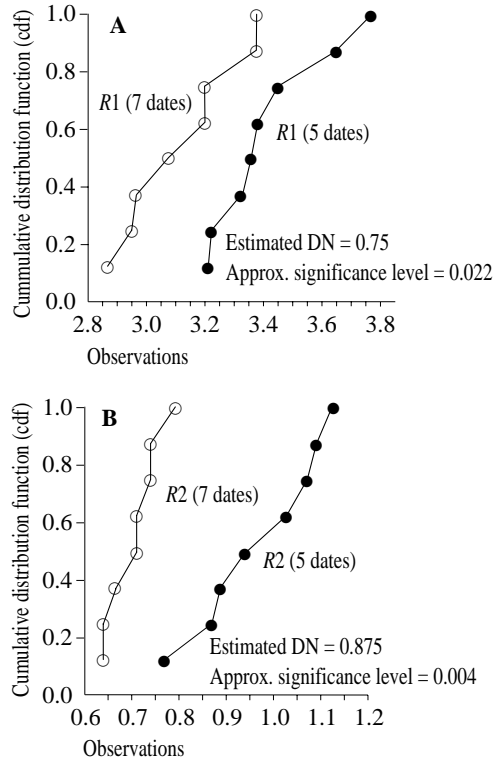


Figure 2. The estimated maximum absolute deviation (DN) between the two cumulative distribution functions (cdf) of 7 and 5 dates of sampling for (A) R1 index (B) R2 index

and night abundances, but significantly influenced richness indices (except R2) (Table 5).

Species were generally abundant at night in both planting patterns. This finding supports results of studies of paddy ecosystem at other locations in Malaysia (Hassan and Ibrahim 1996; Hassan and Yusof 1988). However, the species abundance in the transplanted fields was higher than those of the direct-seeded. This result supports our earlier studies in another rice ecosystem in Malaysia (Hassan and Rashid 1997). The occurrence of higher numbers of insect individuals and species of pests and predators at night contributes to the higher night abundance (Hassan and Rashid 1997). In general, there were no significant differences between direct-seeded and transplanted paddy, with respect to day-

Table 4. Mean of diversity indices for date and time of sampling, F-value statistics and  $\chi^2$  of Kruskal-Wallis test of the significance of diversity indices with the planting type (df = 1) as the main effect. Indices were calculated based on different dates and different times of sampling (n = 7 and n = 8, respectively) in both types of planting

Indices	Indices calculated through					
	Date			Time		
	Mean	F value	$\chi^2$	Mean	F value	$\chi^2$
N	1 231.64	<1.00 <sup>ns</sup>	<1.00 <sup>ns</sup>	1 077.68	<1.00 <sup>ns</sup>	<1.00 <sup>ns</sup>
Richness						
N0	19.64	16.59**	5.79*	20.32	41.11**	11.49**
R1	2.67	12.28**	5.00*	2.77	78.23**	11.29**
R2	0.61	4.69 <sup>ns</sup>	3.92*	0.62	44.19**	11.29**
Diversity						
$\lambda$	0.22	6.23*	6.20*	0.20	24.74**	11.29**
H'	1.96	16.28**	6.20*	1.98	27.04**	11.29**
N1	7.54	17.61**	6.20*	7.48	18.98**	11.29**
N2	5.22	11.38**	6.20*	5.35	13.99**	11.29**
Evenness						
E1	0.66	8.88*	6.20*	0.66	14.84**	10.60**
E2	0.38	5.48*	5.58*	0.36	6.50*	3.98*
E3	0.34	6.69*	5.58*	0.33	8.17*	6.35*
E4	0.70	1.53 <sup>ns</sup>	1.18 <sup>ns</sup>	0.72	3.58 <sup>ns</sup>	2.82 <sup>ns</sup>
E5	0.64	<1.00 <sup>ns</sup>	1.14 <sup>ns</sup>	0.66	<1.00 <sup>ns</sup>	<1.00 <sup>ns</sup>

\* and \*\*significant at 5% and 1% levels respectively  
<sup>ns</sup>not significant

night biodiversity indices, except in species richness indices, N0 and R1. When data were combined by pooling the planting types as an entire ecosystem, the indices values were intermediate between direct-seeded and transplanted (Table 6). The species richness indices, R1 and R2, decreased in direct-seeded treatment because of the lower number of species (N0). Thus, this study indicates that all the richness and diversity indices increase in transplanted rice, with the exception of Simpson's Index ( $\lambda$ ) which decreases with increasing diversity (Table 6). Examination of many definitions of biodiversity indices enables critical evaluation of variable outcome in assessing diversity.

The major advantages of using this SAS programme are that large amounts of raw data sets can be transferred directly into the SAS programme structure through INFILE statement, and the user is free to select or drop the variable/parameter by

using the DROP option. Furthermore, the programme enables simultaneous analyses involving many variables, each with various levels of parameters recorded. Consequently, diversity indices for each parameter observed can be analysed simultaneously. Hence, statistical testing procedure of SAS such as PROC GLM for an analysis of variance (parametric test) or PROC NPAR1WAY (non-parametric test) for each diversity index can be performed. This can be followed by a mean comparison test contrasting the major parameters as sources of variability. Finally, the diversity indices for the entire ecosystem can be calculated.

The comprehensive programme presented here can greatly assist the ecologist to compute the biodiversity indices, followed by subsequent significant comparison testing, hence expediting biodiversity analyses especially when dealing with large data sets. The flexibility to explore effects of ecosystem variables by

Table 5. The diversity indices at day and night arthropod populations of direct-seeded and transplanted paddy. The indices were calculated based on 7 dates of sampling for each planting type. The F-value and  $\chi^2$  of Kruskal-Wallis tests the significance of diversity indices for planting type (df = 1)

Indices	Direct Seeded		Transplanted		F-value	$\chi^2$
	Day	Night	Day	Night		
N	3 929	4 654	3 421	5 242	<1.00 <sup>ns</sup>	< 1.00 <sup>ns</sup>
Richness						
N0	20	20	25	26	121.00**	2.67 <sup>ns</sup>
R1	2.30	2.25	2.95	2.92	573.41**	2.40 <sup>ns</sup>
R2	0.32	0.29	0.43	0.36	5.68 <sup>ns</sup>	2.40 <sup>ns</sup>
Diversity						
$\lambda$	0.23	0.23	0.19	0.12	6.13 <sup>ns</sup>	2.40 <sup>ns</sup>
H'	1.77	1.82	2.05	2.41	5.64 <sup>ns</sup>	2.40 <sup>ns</sup>
N1	5.86	6.16	7.76	11.18	4.05 <sup>ns</sup>	2.40 <sup>ns</sup>
N2	4.26	4.40	5.34	7.80	3.33 <sup>ns</sup>	2.40 <sup>ns</sup>
Evenness						
E1	0.59	0.61	0.64	0.74	2.90 <sup>ns</sup>	2.40 <sup>ns</sup>
E2	0.29	0.31	0.31	0.43	1.33 <sup>ns</sup>	2.40 <sup>ns</sup>
E3	0.26	0.27	0.28	0.41	1.62 <sup>ns</sup>	2.40 <sup>ns</sup>
E4	0.73	0.71	0.69	0.70	11.42 <sup>ns</sup>	2.40 <sup>ns</sup>
E5	0.67	0.66	0.64	0.67	<1.00 <sup>ns</sup>	<1.00 <sup>ns</sup>

\* and \*\*significant at 5% and 1% level respectively

<sup>ns</sup>not significant

addition/elimination process through this programme, enhances critical analysis of changing dynamics of biodiversities.

## References

- Alatalo, R. V. (1981). Problems in the measurement of evenness in ecology. *Oikos* **37**: 199–204
- DeJong, T. M. (1975). A comparison of three diversity indices based on their components of richness and evenness. *Oikos* **26**: 222–7
- Farina, A. (1993). Bird fauna in the changing agricultural landscape. In *Landscape ecology and agroecosystems* (Bunce, R. G. H., Ryszkowski, L. and Paoletti, M. G. ed.) p. 159–67. Boca Raton, Fl.: Lewis Publishers
- Gomez-Sal, A., Alvarez Martioez, J., Munoz-Yanguas, M. A. and Rebollo, S. (1993). Patterns of change in the agrarian landscape in an area of the Cantabrian Mountains (Spain): Assessment by transition probabilities. In *Landscape ecology and agroecosystems* (Bunce, R. G. H., Ryszkowski, L. and Paoletti, M. G. ed.) p. 141–58. Boca Raton, Fl.: Lewis Publishers
- Hassan, S. T. S. and Yusof, A. M. (1988). Day and night visual counts of arthropods on paddy plants. In *Proc. XVIII Intl. Cong. of Entomology* 3–9 July 1988, Vancouver, B.C., Canada, p. 392.
- Hassan, S. T. S. and Ibrahim, R. C. (1996). Comparing day and night population densities of some arthropods on paddy plants. *Malays. Appl. Biol.* **25**(2): 55–9
- Hassan, S. T. S. and Rashid, M. M. (1997). Biodiversity of arthropods of wet paddy ecosystems in Malaysia. *Malays. Appl. Biol.* **26**(1): 45–53
- Heip, C. (1974). A new index measuring evenness. *J. Marine Biol. Assoc.* **54**: 555–7
- Hill, M. O. (1973). Diversity and evenness: a unifying notation and its consequences. *Ecology* **54**: 427–32
- Hurlbert, S. H. (1971). The non-concept of species diversity: A critique and alternative parameters. *Ecology* **52**: 577–86
- Jizhong, Z., Shijun, M. and Changming, C. (1991). An index of ecosystem diversity. *Ecol. Modelling* **59**: 151–63
- Livingston, R. J. (1976). Diurnal and seasonal fluctuations of organisms in a north Florida estuary. *Estuarine and Coastal Marine Sci.* **4**: 373–400
- Lotus 1-2-3 (1993). *Lotus 1-2-3 Release 4 for windows*, Copyright 1991, 1993 Lotus Development Corporation



Table 6. Diversity indices calculated from Tanjung Karang paddy ecosystem in 1992 season for direct-seeded and transplanted rice, with different combination of sampling occasions (date)

No. of sampling occasions (date)	Source of data type	No. of individuals (N)	Diversity Indices											
			Richness					Evenness						
			N0	R1	R2	$\lambda$	H'	N1	N2	E1	E2	E3	E4	E5
14	Combination	17 243	27	2.66	0.21	0.18	2.11	8.26	5.64	0.64	0.31	0.28	0.68	0.64
7	Direct seeded	8 580	21	2.20	0.22	0.22	1.80	6.08	4.36	0.59	0.28	0.25	0.72	0.66
7	Transplanted	8 663	26	2.76	0.28	0.14	2.31	10.09	6.88	0.71	0.38	0.36	0.68	0.64

Ludwig, J. A. and Reynolds, J. F. (1988). *Statistical ecology: A primer on methods and computing* 337 p. New York: John Wiley & Sons

MacIntosh, R. P. (1967). An index of diversity and the relation of certain concepts to diversity. *Ecology* **48**: 392–404

Margalef, R. (1968). *Perspective in ecological theory*. Chicago, IL: University of Chicago Press

Menhinick, E. F. (1964). A comparison of some species-individuals diversity indices applied to samples of field insects. *Ecology* **45**: 859–61

Odum, H. T., Cantlon, J. E. and Kornicker, L. S. (1960). An organizational hierarchy postulate for the interpretation of species-individuals distribution, species entropy, eco-system evolution, and the meaning of a species-variety index. *Ecology* **41**: 395–9

Peet, R. K. (1974). The measurement of species diversity. *Ann. Rev. Ecol. Systems* **5**: 285–307

Pielou, E. C. (1969). *An introduction to mathematical ecology* New York: Wiley-Interscience

Pielou, E. C. (1975). *Ecological Diversity* 165 p. New York: Wiley

SAS Institute (1990). *SAS User's guide, Version 6.03*. Cary, North Carolina: Statistical Analysis System Institute

Shannon, C. E. and Weaver, W. (1949). *The mathematical theory of communication* 117 p. Urbana, IL: University of Illinois Press

Sheldon, A. L. (1969). Equitability indices: dependence on the species count. *Ecology* **50**: 466–7

Simpson, E. H. (1949). Measurement of diversity. *Nature* **163**: 688

Statgraphics (1988). *Statgraphics: Statistical graphics system by statistical graphics corporation*. Maryland: STSC, Inc.

Appendix 1. SAS computer programme for computing index of diversity on fish catch data of Livingston (1976). Data inputs were followed as listed in Table 8.3 of Ludwig and Reynolds (1988)

```

/*          SAS PROCEDURE FOR DIVERSITY INDICES          */
/*          IN STATISTICAL ECOLOGY                      */
/*          e.g.: FISH IN A FLORIDA ESTUARY (LIVINGSTON 1976) */
/*          SAS PROGRAM WRITTEN BY RASHID, M.M.         */

/* All the diversity indices calculated here are well documented by Ludwig and      */
/* Reynolds (1988). Statistical Ecology. A primer on methods and computing.        */
/* John Wiley & Sons, New York. pp. 85–103.                                       */
/* ----- */
/* Formula          Emphasis of Index      References                               */
/* R1=(s-1)/(log(n))      Richness Index   Margalef (1968)                          */
/* R2=(s)/(n**(1/2))      Richness Index   Menhinik (1964)                          */
/* N0=s                   Richness Index   Hill (1973)                               */
/* LAMDA=(ni**2-ni)/(n**2-n)  Diversity Index   Simpson (1949)                           */
/* H'=-((ni/n)*log(ni/n))  Diversity Index   Shannon and Weaver (1949)                */
/* N2=1/lamda             Diversity Index   Hill (1973)                               */
/* N1=exp(h)              Diversity Index   Hill (1973)                               */
/* E1=log(n1)/log(s)      Evenness Index   Pielou (1975)                             */
/* E2=n1/s                Evenness Index   Sheldon (1969)                            */
/* E3=(n1-1)/(s-1)        Evenness Index   Heip (1974)                               */
/* E4=n2/n1                Evenness Index   Hill (1973)                               */
/* E5=(n2-1)/(n1-1)        Evenness Index   Alatalo (1981)                            */
/* d      December catch */
/* j      June catch     */
/* n      total number of individuals */
/* s      number of species (Species richness) */
/* dec   the number of individuals of the ith species December fish collections */
/* june  the number of individuals of the ith species June fish collections */

title 'STATISTICAL ECOLOGY';
data fish;
infile 'a:\fishdata';
input species dec june;
data s;
set fish;
if dec gt 0 then dec=1;
if june gt 0 then june=1;
proc sort; by species;
proc means noprint;
var dec june;
output out=s sum=sd sj;
data n;
set fish;
proc sort; by species;
proc means noprint;

```

```

var dec june;
output out=n sum=nd nj;
data nospp;
set fish;
if dec ge 0 then dec=1;
proc sort; by species;
proc means noprint;
var dec;
output out=nospp sum=nospp;

/* create a new data set */
data nn;
merge nospp n;
do species=1 to nospp;
nd=nd;
nj=nj;
output;
end;

proc sort data=fish;
by species;
proc sort data=nn;
by species;
data Hill;
merge fish nn;
by species;
dlamda1=dec**2-dec;
jlamda1=june**2-june;
hd=-((dec/nd)*log(dec/nd));
hj=-((june/nj)*log(june/nj));
proc means noprint;
var dlamda1 jlamda1 hd hj;
output out=hillShan sum=dlamda1 jlamda1 hd hj;
data BIODIVER;
merge hillShan s n;
R1d=(sd-1)/(log(nd));
R1j=(sj-1)/(log(nj));
R2d=(sd)/(nd**(1/2));
R2j=(sj)/(nj**(1/2));
N0d=sd;
N0j=sj;
dlamda=dlamda1/(nd**2-nd);
jlamda=jlamda1/(nj**2-nj);
n2d=1/dlamda;
n2j=1/jlamda;
n1d=exp(hd);
n1j=exp(hj);
e1d=log(n1d)/log(sd);
e1j=log(n1j)/log(sj);

```

Diversity indices using SAS computer program

```
e2d=n1d/sd;  
e2j=n1j/sj;  
e3d=(n1d-1)/(sd-1);  
e3j=(n1j-1)/(sj-1);  
e4d=n2d/n1d;  
e4j=n2j/n1j;  
e5d=(n2d-1)/(n1d-1);  
e5j=(n2j-1)/(n1j-1);  
title2 'Diversity indices for December collections';  
proc print data=BIODIVER noobs;  
  var nd n0d r1d r2d  
  dlamda hd n1d n2d  
  e1d e2d e3d e4d e5d;  
run;  
title2 'Diversity indices for June collections';  
proc print data=BIODIVER noobs;  
  var nj n0j r1j r2j  
  jlamda hj n1j n2j  
  e1j e2j e3j e4j e5j;  
run;
```

## Appendix 2. SAS computer programme structure in a general form

```

A title 'BIODIVERSITY INDICES ANALYSIS';
data DIVER;
  infile 'd:\data\filename1' eof=next1;
  input type $ date time sppcode hillno num @@;
drop hillno;
return;
  next1: infile 'd:\data\filename2' eof=next2;
  input type $ date time sppcode hillno num @@;
drop hillno;
return;
  next2: infile 'd:\data\filename3' eof=last;
  input type $ date time sppcode hillno num @@;
drop hillno;
return;
  last: infile 'd:\data\filename4';
  input type $ date time sppcode hillno num @@;
drop hillno;
proc sort data=DIVER; by type date time sppcode;

```

```


B title2 'Diversity indices calculated through DATE';
data DFIRST;
  set DIVER;
drop time;
proc sort; by type date sppcode;
proc means noprint;
  var num;
  by type date sppcode;
  output out=sum sum=ni;
proc sort data=sum;
by type date sppcode;
proc print data=sum;
  var type date sppcode ni;
data n;
  set sum;
proc sort; by type date sppcode;
proc means noprint;
  var ni;
  by type date;
output out=n sum=n;
proc sort data=sum;
  by type date sppcode;
data s;
  set sum;
if ni gt 0 then ni=1;
proc sort; by type date sppcode;
proc means noprint;
  var ni;

```

```

    by type date;
output out=s sum=s;
data Hill;
    merge sum n;
    by type date;
    lamda1=ni**2-ni;
    h=-((ni/n)*log(ni/n));
proc means noprint;
    var lamda1 h;
    by type date;
    output out=HillShan sum=lamda1 h;
proc sort data=n;
    by type date;
proc sort data=s;
    by type date;
proc sort data=HillShan;
    by type date;
data BIODIVER;
    merge n s HillShan;
    by type date;
    R1=(s-1)/(log(n));
    R2=(s)/(n**(1/2));
    N0=s;
    lamda=lamda1/(n**2-n);
    n2=1/lamda;
    n1=exp(h);
    e1=log(n1)/log(s);
    e2=n1/s;
    e3=(n1-1)/(s-1);
    e4=n2/n1;
    e5=(n2-1)/(n1-1);
proc sort; by type date;
proc npar1way;
    class type date;
    var n0 r1 r2 lamda h n1 n2 e1 e2 e3 e4 e5
run;
proc glm;
    class type date;
    model n0 r1 r2 lamda h n1 n2 e1 e2 e3 e4 e5=type date;
    means type date / regwq; /* REGWQ      : Equal sample size */
proc print noobs;
    var type date n n0 r1 r2 lamda h n1 n2 e1 e2 e3 e4 e5;
run;

```

 title 'THE WHOLE SEASON';  
 title2 'Diversity indices calculated on all data points combination';  
 data WHOLESYS;  
 set DIVER;  
 drop type date time;

```

proc sort; by sppcode;
proc means noprint;
  var num;
  by sppcode;
  output out=total sum=ni;
proc sort data=total;
  by sppcode;
proc print data=total;
  var sppcode ni;
data n;
  set total;
proc sort; by sppcode;
proc means noprint;
  var ni;
output out=n sum=n;
proc sort data=total;
  by sppcode;
data s;
  set total;
if ni gt 0 then ni=1;
proc sort; by sppcode;
proc means noprint;
  var ni;
output out=s sum=s;
proc sort data=total;
  by sppcode;
data nospp;
  set total;
if ni ge 0 then ni=1;
proc sort; by sppcode;
proc means noprint;
  var ni;
output out=nospp sum=nospp;

  /* create a new data set */

data newn;
  merge nospp n;
  do species=1 to nospp;
  n=n;
output;
end;
proc sort data=total;
  by species;
proc sort data=newn;
  by sppcode;
data HillShan;
  merge total newn;
  by sppcode;

```

Diversity indices using SAS computer program

```
lamda1=ni**2-ni;
H=-((ni/n)*log(ni/n));
proc means noprint;
var lamda1 H;
output out=HillShan sum=lamda1 H;
data BIODIVER;
merge HillShan s n;
R1=(s-1)/(log(n));
R2=(s)/(n**(1/2));
N0=s;
LAMDA=lamda1/(n**2-n);
N2=1/lamda;
N1=exp(h);
E1=log(n1)/log(s);
E2=n1/s;
E3=(n1-1)/(s-1);
E4=n2/n1;
E5=(n2-1)/(n1-1);
proc print noobs;
var n N0 R1 R2 lamda h n1 n2 e1 e2 e3 e4 e5;
run;
```