

A STATISTICAL ANALYSIS OF THE ANGIOSPERM VEGETATION IN MILLER-ANDERSON WOODS NATURE PRESERVE IN ILLINOIS

David L. Quigg
Department of Mathematics
Bradley University
Peoria, IL 61625

Herbert L. Monoson
Department of Biology
Bradley University
Peoria, IL 61625

Caroline F. Schertz
Science Department
Bergan High School
Peoria, IL 61614

ABSTRACT

Statistical methods were used to evaluate the relative abundance of 277 angiosperm species growing in 15 sample plots on nine natural habitats of Miller-Anderson Woods Nature Preserve in Illinois. An ordinal log-linear model was selected to explain the data. Habitat affected the number of observed species and their relative abundance. Predictions of the ordinal model in regard to plant taxa relative abundance were the best for the wooded lowlands, northeast facing slopes, and dry old field succession habitats.

INTRODUCTION

Miller-Anderson Woods Nature Preserve contains 320 acres located west of Illinois Route 29 between the junction of Putnam and Bureau Counties (SW 1/4,

Sec. 36, T15N, R9E, 4th P.M. and part of W 1/4, Sec. 6, T14N, R10E, 4th P.M.). The original land parcel was acquired by the Illinois Department of Conservation (IDOC) in 1969 and consisted of 258 acres. Subsequent land purchases by the IDOC have resulted in the current Nature Preserve acreage.

The Nature Preserve was surveyed as part of the Illinois Natural Areas Inventory (unpublished excerpt) by the IDOC in 1977. The survey team identified and described five natural habitats within the Nature Preserve: dry mesic upland forest; mesic upland forest; sedge meadow; successional field; and glacial hill prairie. The survey team described the Nature Preserve as containing old growth uplands, ravines, and valley forests of the Grand Prairie Section of the Grand Prairie Natural Division of Illinois (Schwegman, 1973).

Monoson and Schertz (1985) divided Miller-Anderson Woods into ten natural habitats on the basis of topographic features (Fig. 1). The ten habitats consisted of dry mesic upland forest, south- and southeast facing wooded slopes, wooded lowlands, north-, northeast, and east-facing wooded slopes, dry old field succession, mesic regrowth forest along stream banks, wet old field succession, sedge meadow, regrowth area along new Illinois Route 29, and back water bog area. A hill prairie located in the dry mesic upland forest was given habitat status because of its unique vegetational composition. Fifteen permanent square 100 m² plots were placed randomly on the first eight habitat designations and the hill prairie for sampling purposes; the regrowth area along new Illinois Route 29 and the back water bog area were not sampled. The parking lot was included in the 1985 study and represented a recently disturbed area. Several physical factors contributed to the small number of sample plots that were used in the study such as differences in area size of the various habitats, variable flowering times of the taxa, steep topographical features, and other abiotic factors.

Monoson and Schertz (1985) reported that a total of 72 families, 185 genera, and 277 species of angiosperms occurred within the fifteen sample plots. Species of the Compositae, Graminae, and Leguminosae families were the best represented. A possible correlation between the relative abundance of these flowering taxa and certain habitats was suggested by the data.

The objective of the present study was to determine if a relationship existed between the relative abundance of plants found within a sample plot and a specific habitat. This relationship would be ecologically important because plant distribution is regulated by topography, water drainage patterns, light duration and intensity, and nutrient availability within any defined land area.

METHODS AND MATERIALS

The habitats examined in this study were dry mesic upland forest (DMUF), south- and southeast facing slopes (SEFS), wooded lowlands (WL), north-, northeast, and east-facing wooded slopes (NEFS), dry old field succession (DOFS), mesic regrowth along stream banks (MR), wet old field succession (WOFS), sedge meadow (SM), and hill prairie (HP). Sample plots were distributed randomly on a grid of these nine habitats that traversed the Nature Preserve. Numerical distribution was six plots in the DMUF, two in the adjacent SEFS, and one each in the WL, NEFS, DOFS, MR, WOFS, SM, and HP habitats.

For each of the plant taxa found in the initial study (Monoson and Schertz, 1985) the habitat, plot number, relative abundance, and flowering date(s) if observed were recorded. Relative abundance values were determined using a modification of the Hanson (1934) procedure. Relative abundance values were rare = <12, occasional = 12-25, and abundant = >25 plants/plot. The relative abundance data (Table 1) are species numbers found in the sample plots and were not representative of the total abundance for the different habitats.

Initially, the Pearson chi-square test for independence (Bhattacharyya and Johnson, 1977) was used to determine if a relationship between sample plot and plant relative abundance existed. The relationship of habitat to relative abundance was examined by a second application of the chi-square test. The second test used an abbreviated table of data formed by adding abundance numbers for plots within the same habitat, e.g. southeast facing slopes (SEFS) total 12 taxa in rare, 21 taxa in occasional, and 17 taxa in abundant categories. Dry mesic upland forest (DMUF) plots were combined within either the northern or southern plateau areas to allow for possible geographic effects. This separation of DMUF plots resulted in ten habitat designations for the statistical analysis.

Several statistical models were considered in the assessment of the variability of the data. These were the log-linear models of Fienberg (1977) adapted to use habitat as the row variable. This method allowed the logarithm of the expected number in a cell of Table 1 to depend upon (at least) the habitat and the overall frequency in "rare", "occasional" and "abundant" categories. All of the data contained in Table 1 was used because the sample plot was considered to be the basic experimental unit.

The structure of Table 1 included the order of the relative abundance categories — "rare" is less than "occasional" is less than "abundant". Therefore, a log-linear type of model (Agresti, 1984) that had ordinal column categories was considered. This ordinal model was selected for further study.

The ordinal model is described by the equation

$$\log m_{ij} = \mu + \lambda_i + \alpha_j + \tau_i v_j$$

where the variables are defined as follows:

- m_{ij} = expected number in a cell in column j with habitat i
- μ = overall average of $\log m_{ij}$
- λ_i = the direct effect of the i^{th} habitat on the log expected number
- α_j = the direct effect of the j^{th} column (abundance category) on the log expected number
- τ_i = the interaction effect between habitat i and relative abundance
- v_j = relative abundance score (predetermined) reflecting the ordering of the columns: $v_1 = -1$, $v_2 = 0$, $v_3 = 1$

The equation is subject to the conditions $\sum \lambda_i = 0$, $\sum \alpha_j = 0$, and $\sum \tau_i = 0$. The DMUF plots on the northern and southern plateaus were considered to be different habitats for this analysis.

With a minimum number of additional parameters (the τ_i 's), the ordinal model allowed a statistical consideration of the concept that the habitat influenced both the number of species per plot and the allocation of these species in each relative abundance category. Habitat would affect only total numbers of species per plot and not relative abundance if the τ_i 's were all zero.

Parameter estimation for the model was performed using a FORTRAN program (unpublished) written by the first author. Standard log-linear statistical packages were not appropriate due to both the irregular replication pattern, i.e. variable numbers of plots per habitat, and the ordinal nature of the relative abundance categories. Usually, normal categories are assumed by most log-linear statistical package programs. For each parameter, the partial derivative of the log-likelihood was set to zero, and the resulting system of equations solved numerically using the International Mathematical and Statistical Libraries, Incorporated (IMSL, 1982) subroutine ZSCNT. Subroutine ZSCNT solves a system of nonlinear equations using the secant method. The program was executed on the Bradley University Cyber 830 mainframe computer.

RESULTS AND DISCUSSION

The number of plant species collected from the fifteen plot locations of Miller-Anderson Woods Nature Preserve can be seen in Table 1. Slight differences in the data are noted when compared to the previous published information of Monoson and Schertz (1985). In the present study, these differences are due to a particular taxon occurring in more than one sample plot.

Total plant taxa numbers varied according to the habitat (Table 1). Wooded lowlands and mesic regrowth along stream banks contained the highest numbers of different taxa recorded for the study. Presumably, abiotic and other ecological conditions in the two habitats contributed to the high number of plant species.

The Pearson chi-square (X^2) was 56.06 for the data in Table 1. This statistical test would follow a chi-square distribution with 28 degrees of freedom if sample plot and relative abundance were independent. Independence of the latter two variables was rejected since the X^2 exceeded even a 0.005 critical value of 50.99.

A 10-by-3 table was formed when sample plots formed from the same habitat were combined. The Pearson chi-square test statistic (X^2) was 45.22 for this table. Independence of habitat and relative abundance was rejected because the X^2 exceeded the 0.005 critical value of 37.16 for a chi-square distribution with 18 degrees of freedom.

The ordinal model provided a marginally satisfactory fit of the data. As there were ten habitats, the model used 21 parameters- one μ , nine λ_i 's, two α_j 's, and nine τ_i 's. Parameter estimates were calculated and then the expected cell frequencies (m_{ij} 's) were determined. Expected and observed numbers per cell in the ordinal model can be seen in Table 2.

The goodness-of-fit test statistics, likelihood-ratio chi-square G^2 and Pearson chi-square X^2 , were 38.63 and 40.95 respectively for Table 2. These statistics, under the null hypothesis of satisfactory fit, would have a chi-square distribution of $45-21 = 24$ degrees of freedom. Thus, the ordinal model G^2 and X^2 have significance of approximately 0.035 and 0.02 respectively. While the ordinal model was only marginally acceptable, the alternatives of independence and further parameterization were unacceptable. The log-linear model of next higher complexity would have nine additional habitat parameters and would leave these habitats uniquely characterized by a single sample plot with zero degrees of freedom toward a chi-square test of satisfactory fit.

We also examined the null hypothesis that $\tau_i = 0$ for all i 's in the ordinal model. The likelihood-ratio chi-square test of the null hypothesis yielded a G^2 of 26.44 based on 9 degrees of freedom. The null hypothesis was rejected because G^2 exceeded the 0.005 critical value of 23.59. Therefore, the inclusion of habitat-versus-relative-abundance interaction contributed significantly towards an adequate interpretation of the data.

Maintenance of separate parameters for northern and southern DMUF plateaus aided in the model-fitting process. Associating all of the DMUF plots with a single set of parameters would have added one to the degrees of freedom and more than six to the ordinal model's G^2 .

The ordinal model predicted relative abundance for some habitat sample plots better than for others (Table 2). Observed and expected values were very similar for the WI, NEFS, and DOFS habitat plots. Conversely, the observed and expected values in the MR habitat plot were quite different. The MR plot with only one degree of freedom, out of 24 degrees of freedom in the ordinal model, contributed roughly 25 percent to the chi-square statistics. Exclusion of the MR plot would result in a more acceptable fit of the data to the model. The similarity or dissimilarity of the observed and expected plant relative abundance values were probably linked to ecological considerations in the different habitats. The validity of the different statistical parameters used in this investigation can be tested by their repeatability. At the present time, one can only speculate as to which of the abiotic factors was the main contributor to plant relative abundance in a particular habitat.

In summary, we believe that the ordinal model best explained the data. Habitat affected the number of observed species and their relative abundance tendencies. We think that future studies ideally should include two or more plots per habitat in the experimental design. Having multiple plots per habitat would assist in the evaluation of distribution patterns of species abundance within habitats.

ACKNOWLEDGMENTS

The authors would like to thank the Illinois Nature Preserves Commission (INPC), Illinois Department of Conservation for their permission to conduct this investigation of Miller-Anderson Woods Nature Preserve. This study would not have been possible without the financial support provided by the Board of Research and Creativity of Bradley University. We thank Dr. Barbara Frase for her critical review of the manuscript.

LITERATURE CITED

- Agresti, A. 1984. Analysis of ordinal categorical data. John Wiley & Sons, New York, 287 pp.
- Bhattacharyya, G.K. and R.A. Johnson. 1977. Statistical concepts and methods. John Wiley and Sons, New York, 639 pp.
- Finberg, S.E. 1977. The analysis of cross-classified categorical data. The Massachusetts Institute of Technology Press, Cambridge, 151 pp.
- Haber, M. 1985. Maximum likelihood methods for linear and log-linear models in categorical data. Computational Statistics and Data Analysis, 3:1-10.
- Hanson, H.C. 1934. A comparison of methods of botanical analysis of the native prairie in western North Dakota. Jour. Agricultural Res. 49:815-842.
- International Mathematical and Statistical Libraries, Inc. 1982. IMSL Library Reference Manual. IMSL, Inc., Houston, Texas.
- Monoson, H.L. and C.F. Schertz. 1985. Angiosperm flora of Miller-Anderson Woods Nature Preserve Bureau and Putnam Counties, Illinois. Trans. Ill. Acad. Sci. 78:263-279.
- Schwegman, J. 1973. Comprehensive plan for the Illinois Natures Preserves System. Part 2: The natural divisions of Illinois. Illinois Nature Preserves Commission. Rockford, Illinois.

Table 1. Sample Plot by Frequency Category of Observed Species

Habitat ^a	Abundance			Total
	Rare ^b	Occasional ^c	Abundant ^d	
DMUF-North				
Plot 1	0	8	13	21
Plot 2	4	8	6	18
Plot 6	1	9	13	23
DMUF-South				
Plot 12	2	3	5	10
Plot 13	4	6	9	19
Plot 14	0	3	6	9
SEFS				
Plot 3	7	11	6	24
Plot 5	5	10	11	26
WL - Plot 4	17	43	19	79
NEFS - Plot 11	6	18	12	36
DOFS - Plot 7	3	19	17	39
MR - Plot 8	4	45	18	67
WOFS - Plot 9	1	17	8	26
SM - Plot 10	7	12	7	26
HP - Plot 15	8	19	6	33
Total	69	231	156	456

^aDMUF = Dry mesic upland forest; SEFS = Southeast facing slopes; WL = Wooded lowlands; NEFS = Northeast facing slopes; DOFS = Dry old field succession; MR = Mesic regrowth along stream banks; WOFS = Wet old field succession; SM = Sedge meadow; HP = Hill prairie.

^b<12 plants

^c12-15 plants

^d>25 plants

Table 2. Observed and Expected Relative Abundance Numbers Under the Ordinal Model

Habitat ^a	Rare ^b	Occasional ^c	Abundant ^d
DMUF - North			
Plot 1 observed	0	8	13
expected	1.36	8.94	10.36
Plot 2	4	8	6
expected	1.36	8.94	10.36
Plot 6	1	9	13
expected	1.36	8.94	10.36
DMUF - South			
Plot 12	2	3	5
expected	1.06	5.87	5.73
Plot 13	4	6	9
expected	1.06	5.87	5.73
Plot 14	0	3	6
expected	1.06	5.87	5.73
SEFS			
Plot 3	7	11	6
expected	4.54	13.43	7.04
Plot 5	5	10	11
expected	4.54	13.43	7.04
WL - Plot 4	17	43	19
expected	17.06	42.88	19.06
NEFS - Plot 11	6	18	12
expected	5.52	18.96	11.52
DOFS - Plot 7	3	19	17
expected	3.38	18.23	17.38
MR - Plot 8	4	45	18
expected	9.17	34.66	23.17
WOFS - Plot 9	1	17	8
expected	3.00	13.01	10.00
SM - Plot 10	7	12	7
expected	5.94	14.12	5.94
HP - Plot 15	8	19	6
expected	8.57	17.85	6.57

^aDMUF = Dry mesic upland forest; SEFS = Southeast facing slopes; WL = Wooded lowlands; NEFS = Northeast facing slopes; DOFS = Dry old field succession; MR = Mesic regrowth along stream banks; WOFS = Wet old field succession; SM = Sedge meadow; HP = Hill prairie.

^b<12 plants

^c12-25 plants

^d>25 plants

