COENOSES 8(1): 21-24, 1993

MARKOV MODELS OF SPATIAL DEPENDENCE IN VEGETATION

M. R. T. Dale, G. H. R. Henry and C. Young

Department of Botany, University of Alberta, Edmonton, Alberta, T6G 2E9, Canada

Keywords: Ellesmere Island, Goodness-of-fit, Meadows, Transects,

Abstract. The purpose of this study was to evaluate the spatial dependence of species composition in transects of contiguous quadrats. Three aspects were examined. The main analysis was a set of goodness-of-fit tests to determine whether low order Markov models were adequate for describing species composition, looking both at single species and multiple species combinations. The second complementary method was an examination of the means and variances of runs of absences in the data, as defined relative to single species and to combinations of up to six species. The third analysis evaluated the similarity in composition of quadrats at spacings from one to ten quadrats.

These methods were applied to vascular plant data from transects 1000 quadrats long in sedge meadows on Ellesmere Island. While the goodness of fit tests suggested that low order Markov models provided an adequate description of the many species data, the runs tests and local similarities showed that they were not appropriate.

Introduction

When vegetation is sampled with a string of equidistant points or a transect of contiguous quadrats and only presence or absence is considered, the record of each species can be represented by a sequence of 0's (absences) and 1's (presences). It is tempting to consider treating such data, at least initially, as if the 0's and 1's were the result of a stochastic process of some kind. The stochastic process that seems logical to consider is a two-state Markov process. The two states are, of course, 0 and 1, and the Markov property for a m-th order process is that the state of the i-th stage of the process depends only on the states of the m stages preceding it, not on the entire history of the process. Thus, for example, in a Markov process of order 0, the state at any place in the process is completely independent of what has gone before. On the other hand, in a Markov process of order 1, the state at any place in the process depends only on the state immediately preceding it.

A number of studies have applied Markov models to ecological succession, by using them to described changes in species composition through time (*e.g.* Horn 1975, van Hulst 1979). Binkley (1980) and Lippe *et al.* (1985) have tested the adequacy of such models and found that they are not good descriptions of temporal changes, chiefly because the transition probabilities themselves are not constant through time. In this paper, we examine the application of such models not to temporal change but to spatial changes in species composition along transects of quadrats. These data are expected to exhibit a degree of spatial dependence (or autocorrelation) because the presence or absence of species in a quadrat may not be completely independent of presence or absence in adjacent quadrats.

This spatial dependence is related to the fact that vegetation, in general, exhibits spatial pattern, which is nonrandomness of a certain periodicity, and frequently each species can have pattern at several scales (Dale and Blundon 1990). This means that when single species are considered, there is no reason to expect that any will be well modelled by a low order Markov process. However, it is conceivable that when several species (say k of them) are considered at once, a low order Markov model (now with 2^k states) might be appropriate. On the other hand, if there is multiple species pattern consisting of distinct phases each with consistant composition, low order models will continue to be inappropriate no matter how many species are considered. Even if low order Markov models cannot be rejected, other aspects of the vegetation should also be considered, for example, the means and variances of the lengths of runs of 1's and 0's in the data. Pielou (1977) comments that in studies of *pairs* of plant species in natural vegetation the variation in run length was always greater than expected and that, based on her studies, it seems unlikely that a first order Markov model will often be appropriate for them. A third kind of evidence to consider is the similarity in species composition of quadrats as a function of their distance from each other.

In this paper we examine vegetation data from the high arctic for these characteristics, using single species and multiple species combinations. We test whether low order Markov models are adequate descriptions of spatial dependence in the data, we look at the means

Table 1. Species list and numbers of quadrats in which each was present. A blank indicates less than 25 quadrats. The mean and variance of the number of *these* species per quadrat are given.

	Transect			
Species	BMS	BRS	OWT	CRT
1. *Carex aquatilis* Wahleb.	417	391	205	895
2. *C. membranacea* Hook.			485	
3. *C. misandra* R.Br.			150	
4. *Eriophorum triste* (Th.Fr.) Hadac & Löve	899	900	866	264
5. *E. scheurchzeri* Hoppe	521	457		145
6. *Juncus biglumis* L.	396	166	57	29
7. *Dupontia fisheri* R. Br.				881
8. *Arctagrostis latifolia* (R.Br.) Griseb.	229	422		71
9. *Pleuropogon sabinei* R.Br.	150	81		
10. *Salix arctica* Pall.	319	310	93	56
11. *Dryas integrifolia* M. Vahl			441	
12. *Polygonum viviparum* L.	97	76	195	31
13. *Saxifraga hirculus* L.				39
14. *S. cernua* L.				256
15. *S. oppositifolia* L.			147	
16. *Pedicularis capitata* Adams			27	
17. *Draba lactea* Adams			30	
18. *Cardamine pratensis* L.				61
19. *Stellaria longipes* Goldie				32
20. *Equisetum variegatum* Schleich			26	
means	2.80	3.03	2.74	2.76
variance	0.91	1.05	1.35	0.90

and variances of the lengths of runs of zeros, and we test local similarity.

Methods

Field Methods

The data consist of presence-absence records in four 100m transects, each consisting of 1000 10cm x 5cm quadrats. Three of the transects are from Sverdrup Pass, Ellesmere Island (79° 09'N, 79° 40'W), two of which were in the same meadow, layed out at right angles to each other (designated BMS and BRS), and the third in different meadow (CRT). The fourth transect is from Alexandra Fiord, Ellesmere Island (78° 53'N,75° 55'W) and is designated OWT. The vegetation was primarily sedge meadow with pronounced hummock-hollow physiognomy only in the last transect. For the purposes of analysis, we considered only those vascular species that were present in 25 or more of the quadrats in a transect. There were therefore eight species used in BMS and BRS and twelve in CRT and OWT (see Table 1).

Analysis

The first kind of test performed was to examine whether the data fit Markov processes of a given order. These take the form of goodness-of-fit tests (cf. Anderson and Goodman 1957, Kullback et al. 1962), with the test statistic being compared to the χ^2 distribution with the appropriate degrees of freedom:

Order 0:

$$G = 2 \sum \sum f_{ij} \log (f_{ij}\ f_{..} / f_{i.} f_{.j}) \qquad d.f. = (r-1)^2$$

where the f's are observed frequencies with the dot notation indicating summation over the dotted variable, and r is the number of states in the process.

Order 1:

$$G = 2 \sum \sum \sum f_{ijk} \log (f_{ijk}\ f_{.j.} / f_{ij.} f_{.jk}) \qquad d.f. = r(r-1)^2$$

Order 2:

$$G = 2 \sum \sum \sum \sum f_{ijkl} \log (f_{ijkl}\ f_{.jk.} / f_{ijk.}\ f_{.jkl}) \quad d.f. = r^2 (r-1)^2$$

These three tests were carried out first on all species individually, then on all pairs of species, and lastly on all triplets of species. Tests of order 1 were carried out on groupings of up to four species, and tests of order 0 up to six species. We recognise that within each set, the tests are not all independent, but this lack of independence does not affect this study since we do not seek to base conclusions on any particular test. Increasing the order of the model tested or the number of species considered increases the number of degrees of freedom and thereby decreases the expected values causing problems for the reliability of the tests (*cf.* van Hulst 1979).

The second analysis that was performed was to look at the lengths of the runs of 0's for each species. If adjacent quadrats are independent (a zero order Markov process), and if the proportion of 1's is p_1, then the expected length of runs of 0's is $1/p_1 = \mu$ with a variance of $\mu^2 - \mu$ (Pielou 1977). We examined runs of 0's for all single species, for species pairs and for combinations of up to six species. Since we are not interested in the individual species or sets of them in particular, but only in general trends, we did not look for significance in each analysis, but only in the overall study using the sign test.

The third analysis was to test the "local similarity" within the transects. The average similarity of species composition of quadrats at spacings of 1,2,...10 were calculated using the simple matching coefficient. (The distance of ten quadrats was chosen somewhat arbitrarily, but all four transects had species with scales of less than 8, as determined by Three Term Local Quadrat Variance (Hill 1973)). The transects were then "shuffled" (*i.e.* reordered randomly by exchanging pairs of quadrats) 100 times with the average similarity recalculated after each shuffle. This procedure enabled us to determine whether the observed local similarities

Table 2. Proportions of significant results for the Markov goodness-of-fit tests.

	Number of species					
Transect	1	2	3	4	5	6
Order 0						
BMS	all	all	all	all	55/56	none
BRS	all	all	all	all	41/56	none
OWT	all	all	219/220	414/495	none	none
CRT	all	all	214/220	285/495	none	none
Order 1						
BMS	7/8	27/28	7/56	none		
BRS	8/8	27/28	5/56	none		
OWT	5/12	18/66	none	none		
CRT	8/12	28/66	none	none		
Order 2						
BMS	4/8	7/28	none			
BRS	6/8	8/28	none			
OWT	3/12	2/66	none			
CRT	8/12	1/66	none			

were greater than expected from the composition of the quadrats.

Results

The Markov goodness-of-fit tests showed that as higher order models or more species were considered, fewer and fewer of the tests gave a significant result (Table 2). All tests related to models of zero order were significant for single species or pairs, but none were significant when the combinations of six species were considered.

In the runs tests for single species and for combinations of species from two to six, the average run length was greater than expected in almost all cases (the exceptions being two single species and one pair of species in transect CRT). The variance of run length was greater than expected in all cases (cf. Table 3 for single species). The ratio of observed to expected was always greater for the standard deviation than for the mean, indicating that the increase in variance was not just a result of increased mean run length. The averages of these ratios are given in Table 4. The near universality of these characteristics causes them to be statistically significant.

The randomization tests of local similarity showed that the observed values were significantly great in all cases for spacings from 1 to 10 quadrats. The observed values are given in Table 5.

Discussion

Even when there is strong spatial dependence within each species (spatial pattern), as the number of species considered increases, the probability that some change occurs between quadrat i and quadrat i+1 increases. Thus the apparent degree of a Markov process modelling the the vegetation decreases as more species are considered. This is one way of explaining the observation in the spatial autocorrelation apparent in contingency table testing of contiguous quadrat data decreases with increasing k (Dale *et al.* 1991). The success of a low order Markov model in describing vegetation might seem to suggest that there can be no multispecies pattern in that vegetation. However, that

Table 3. Results of testing the lengths of runs of zeros using single species.

Transect		Observed/expected	
Species	No. of runs	Mean	s.d.
BMS			
1	43	5.57	14.09
4	57	1.59	4.14
5	149	1.68	2.37
6	179	1.33	2.09
8	60	2.90	11.30
9	81	1.57	2.53
10	105	2.06	3.36
12	51	1.67	2.61
BRS			
1	57	4.18	7.33
4	59	1.52	4.58
5	147	1.68	3.14
6	102	1.35	2.04
8	122	2.00	4.98
9	52	1.13	1.87
10	125	1.69	2.67
12	48	1.46	2.50
CRT			
1	76	1.24	2.43
4	101	1.92	4.75
5	98	1.25	2.46
6	25	0.89	1.31
7	77	1.36	2.91
8	49	1.28	2.48
10	31	1.58	3.15
12	24	0.89	1.72
13	25	1.49	2.79
14	122	1.55	3.85
18	42	1.33	5.40
19	30	1.03	2.11
OWT			
1	116	1.38	2.36
2	177	1.41	1.96
3	100	1.26	2.15
4	96	1.20	1.81
6	40	1.34	1.46
10	53	1.41	1.58
11	138	1.77	2.46
12	115	1.34	1.74
15	65	1.92	1.90
16	14	1.73	1.87
17	24	1.17	1.49
20	31	1.10	2.53

Table 4. The average ratios of observed to expected mean and observed to expected standard deviations of the length of runs of zeros when 1, 2,...,6 species are considered.

Transect	No. of species	Means	St. deviations
BMS	1	2.30	5.31
	2	1.89	3.95
	3	1.64	3.91
	4	1.47	4.02
	5	1.35	4.06
	6	1.26	3.79
BRS	1	1.88	3.64
	2	1.75	3.70
	3	1.60	3.87
	4	1.47	4.06
	5	1.36	3.98
	6	1.26	3.67
OWT	1	1.42	1.94
	2	1.42	1.95
	3	1.39	1.98
	4	1.35	2.01
	5	1.31	2.04
	6	1.27	2.08
CRT	1	1.32	2.95
	2	1.38	2.98
	3	1.36	3.01
	4	1.33	3.03
	5	1.30	3.03
	6	1.27	3.04

is not true. In the Markov model, all changes of state are regarded as equivalent. In real vegetation, quadrats close to each other may have nonidentical species compsitions and yet be similar. A single species may change from presence to absence between quadrat i and quadrat i+1 and then change back to presence between i+1 and i+2. Thus there may be runs of quadrats that are similar to each other, but not identical, as our tests of local similarity illustrate.

As noted above, higher order tests or tests involving more species lower the expected values in the tests. Small expected values produce a bias toward incorrectly rejecting the null hypothesis (Zar 1984, p.49), just the opposite of the trend in our tests.

In comparing the results for the four transects, there seems to be a clear difference between BMS and BRS on the one hand and OWT and CRT on the other. This is particularly obvious in the results of tests of 0 order models (Table 2; the latter pair of transects have the test result become non-significant with fewer species. We believe that the explanation for this fact lies in the fact that all four transects have an average of about three of the common species per quadrat, but since the latter transects draw from a pool of twelve, rather than only eight, spatial dependence is less evident.

When interpreting the results presented here, it is important to realise that the fact that a test is not significant does not prove that the process is indeed a

Table 5. Local similarity of quadrats at distances of 1 to 10 quadrat lengths. All values are significantly high, based on randomization tests.

	Transect			
Distance	BMS	BRS	OWT	CRT
1	.818	.822	.838	.883
2	.791	.799	.802	.868
3	.773	.794	.793	.865
4	.765	.787	.789	.862
5	.755	.781	.782	.860
6	.759	.771	.779	.859
7	.759	.760	.781	.861
8	.750	.757	.777	.861
9	.744	.754	.778	.856
10	.744	.750	.774	.857

Markov process of a particular order. Therefore, while the goodness of fit tests provide no basis for rejecting a low order Markov model as a description of multiple species spatial dependence, the tests of run lengths (means and variances) and of local similarity show that a zero order Markov model is not an appropriate description.

Acknowledgements. This research was supported by the Natural Sciences and Engineering Research Council of Canada and the Canadian Circumpolar Institute.

REFERENCES

Anderson, T. W. & L. A. Goodman. 1957. Statistical inference about Markov chains. Ann. Math. Stat. 28:89-109.

Binkley, C. S. 1980. Is succession in hardwood forests a stationary Markov process? For. Sci. 26:566-570.

Dale, M. R. T. & D. J. Blundon. 1990. Quadrat variance and pattern development during primary succession. J. Veg. Sci. 1:153-164.

Dale, M. R. T. , D. J. Blundon, D. A. MacIsaac & A. G. Thomas. 1991. Multiple species effects and spatial autocorrelation in detecting species associations. J. Veg. Sci. 2:635-642.

Hill, M. O. 1973. The intensity of spatial pattern in plant communities. J. Ecol. 49:225-235.

Horn, H. S. 1975. Markovian properties of forest succession. In: Cody, M. L. & M. J. Diamond (eds). *Ecology and Evolution of Communities.* Harvard U. Press, Cambridge MA.

Hulst, R. van, 1979. On the dynamics of vegetation: Markov chains as models of succession. Vegetatio 40:3-14.

Kullback,S. , M. Kupperman & H. H. Ku. 1962. Tests for contingency tables and Markov chains. Technometrics 4:572-608.

Lippe, E. , J. T. de Smidt, & D. C. Glenn-Lewin. 1985. Markov models of succession: a test from a heathland in the Netherlands. J. Ecol. 73:775-791.

Pielou, E. C. 1977. *Mathematical Ecology.* Wiley, New York.

Porsild, A. E. & W. J. Cody. 1980. *Vascular Plants of Continental Northwest Territories, Canada.* National Museums of Canada, Ottawa.

Zar, J. 1984. *Biostatistical Analysis.* Prentice Hall, Englewood Cliffs, N. J.

Manuscript received: June 1992