COENOSES 10(1): 45-49, 1995

A MONTE CARLO RANDOMISATION PROCEDURE FOR ESTIMATING THE SYNCHRONITY OF GROWTH-CURVES IN PLANT-DEMOGRAPHY STUDIES

J. Geißelbrecht, L. Mucina* and C.K. Hiebinger

Institute of Botany, University of Vienna, Rennweg 14, A-1030 Wien, Austria (* corresponding author)

Keywords: *Anthyllis vulneraria*, *Globularia punctata*, Leaf natality and mortality, Population dynamics, Synchronity of processes

Abstract. A method for comparing growth-curves used to describe the dynamics of individuals or their parts in plant-demographic studies is presented. The shape of the curves is compared using a multiplication invariant resemblance measure. A possibility of shifting vector elements ss also considered. The synchronity of curves is tested using a Monte Carlo randomization procedure yielding many pairs of semirandom test vectors. The test vectors were constructed to be permutations of the original vectors. The power of the method is shown on an example involving cumulative natality and mortality curves for *Globularia punctata* and *Anthyllis vulneraria*, respectively (data collected in a dry grassland in eastern Austria).

Introduction

In population biological studies aimed at the demography of plants or their parts, growth curves depicting the processes of birth (natality) or death (mortality) are often used (Sarukhán & Harper 1973, Sydes 1984, Mitchley 1988, van der Sman *et al.* 1993, Houle 1994, etc.). Depending on the genetic (internal) or environmental (external - incl. climate, nutrient availability, grazing pressure, etc.) variability of individuals, the curves may differ both in qualitative (e.g., shape) and quantitative (e.g., absolute cumulative number of individuals) terms. Concentrating only on shapes of curves, the similarity can be judged either purely intuitively by optical judgement or by application of appropriate statistical tools. In the sequel we attempt to present a simple method (based on comparison of two vectors) to quantify the similarity between two demography curves, both with or without application of temporal shift.

Small number of individuals compared in the above-mentioned demography curves poses a serious statistical problem on generalizations based on such data. Under controlled environmental conditions (exclusion of environmental heterogeneity), the synchronity of the curves (in other words, their similarity in shape) can be a useful aid to overcome the problems of low number of data. To show that the pattern of the curves (their shape) is not random as compared to a null-model, a Monte Carlo randomization procedure was included into the analysis of two-curve synchronity. To develop and show the power of the method we use cumulative curves of dead and young leaves in two dry grassland species observed during one vegetation season.

Structuring the data

A cumulative curve of (dead or newborn) leaves is a (temporally) ordered data set, which can be conceived as an n-dimensional vector **z**, with n+1 being the number of days when an observation was made. It is plausible not to consider **z** for further calculations but a vector **v**, which contains the mean increase during the respective period, obtained by dividing each element by the number of days separating two observation dates. This procedure leads to the construction of a vector containing the full information for each individual using this procedure.

Example 1

observation day	0	4	7	11	14	19	23	30	33	37
leaf number (individual 1)	0	6	8	13	15	16	16	16	21	28
leaf number (individual 2)	0	3	5	8	9	9	10	12	16	19

$\mathbf{z}_1 = (6,2,5,2,1,0,0,5,7)$

$\mathbf{z}_2 = (3,2,3,1,0,1,2,4,3)$

$\mathbf{v}_1 = (1.50,0.67,1.25,0.67,0.20,0.00,0.00,1.67,1.75)$

$\mathbf{v}_2 = (0.75,0.67,0.75,0.33,0.00,0.25,0.29,1.33,0.75)$

Defining resemblance

As we are interested in similarity of shapes of curves rather than of absolute values, it is convenient to define a resemblance measure which is insensitive to the multiplication of one or both vectors with a real number 0 (multiplication invariance). Two vectors shall be considered as perfectly similar if one is a multiple of the other; this implies that the two vectors are parallel. On the other hand they shall be perfectly dissimilar if they have not got (or lost) even one leaf within the same period, speaking in geometric terms - if their inner product is 0 or if they are orthogonal to each other, respectively. These conditions are met by the cosine of the angle between the two vectors, which is calculated as

$$S = \cos(v1,v2) = \langle v1,v2\rangle / [\langle v1,v1\rangle\langle v2,v2\rangle]^{1/2} \qquad (1)$$

or

$$S = \Sigma\, v_1^i v_2^i \,/\, [\,\Sigma\,(v_1^i)^2\,\Sigma\,(v_2^i)^2]^{1/2} \qquad (2)$$

where <.,.> denotes the inner or dot product. S is confined to the interval [0,1] because all elements of the vectors are greater than or equal to 0. As S is in fact the inner product of the normalized vectors, the multiplication invariance is also fulfilled. A further important feature of S is that, at least for random vectors, it is practically independent from their dimension if dim>10, which was found out empirically.

In the above example S becomes

$S = 6.2661 \,/\, [10.6017\ 4.1608]^{1/2} = 6.2661/6.6417 = 0.9434$,

which denotes quite a high similarity. However, it is necessary to specify more quantitatively how "high" this similarity is.

Correcting for time shifts

Consider the vectors $\mathbf{u_1} = (0,1,0,0,2,0,0)$ and $\mathbf{u_2} = (0,0,2,0,0,4,0)$. They have a similarity of 0 although the respective curves would have exactly the same shape if $\mathbf{u_2}$ were shifted one to the left, i.e. if we allow for a temporal shift. In order to detect similarities apart from the temporal shifts, it is convenient to include a possibility of shifting vector elements to the left or right by adding a degree of freedom in the calculation. We loose, however, one vector dimension per shift of one element. Having m vectors to be compared with each other, there are (m-1) degrees of freedom. Since only the relative temporal shifts are of interest, we can fix the first vector and shift each of the others conveniently. Certainly only a limited shift is suggestive, because if there remains only one dimension, all vector combinations lead to S = 1.

Testing for significance

The main problem is to set up a reasonable null hypothesis which is not too far from reality so that a rejection of it is not trivial. A plausible null hypothesis may read: *The similarity between the two vectors is not different from the similarity between two semirandom test vectors. The notion of semi-randomness implies that the test vectors are somehow determined by the original vectors.*

A possibility to construct an n-dimensional test vector is to draw n elements from the original vector at random. Drawing with replacement emphasizes a distributional aspect, i.e. the original vector elements define a distribution from which elements for the test vector are taken. We can construct n^n different vectors in this way if all distribution elements are different. Drawing without replacement emphasizes a sequential aspect, i.e. the original vector is one of n! arrangements of n elements. It depends on the ecological question which aspect is given preference. However, the distributional aspect becomes less meaningful the smaller the vector dimension is. Therefore we concentrate on the sequential aspect here, i.e. all test vectors are permutations of the original vectors. This has a great computational advantage: it is sufficient to permute only one of the compared vectors because the ordering of the second vector does not influence the sums to be calculated by formula (2).

By the construction of many pairs of semirandom test vectors, one can make a randomization test by counting how often there is a greater similarity than between the two original vectors. When we want to show that two vectors are significantly similar, it is appropriate to make a one-sided test. The significance value, i.e. the probability that the decision to reject the null hypothesis was incorrect, is then the number of runs resulting in a higher similarity divided by the total number of runs.

Example 2

(for the data see Example 1)

actual similarity between $\mathbf{v_1}$ and $\mathbf{v_2}$: $S_1 = 0.9434$

Run 1

$\mathbf{v_1}' = (1.50,0.67,1.25,0.00,0.67,0.20,0.00,1.67,1.75)$

$\mathbf{v_2}' = \mathbf{v_2}$

$S_0^{(1)} = 0.9176$

Run 2

$\mathbf{v_1}' = (1.50,0.67,0.00,0.20,1.67,0.00,1.25,1.75,0.67)$

$\mathbf{v_2}' = \mathbf{v_2}$

$S_0^{(2)} = 0.7270$

Run 3

$\mathbf{v_1}' = (1.25,0.20,1.50,0.67,1.67,0.00,0.67,0.00,1.75)$

$\mathbf{v_2}' = \mathbf{v_2}$

$S_0^{(3)} = 0.5906$

Run 4

$\mathbf{v_1}' = (0.67,1.75,1.67,1.25,1.50,0.67,0.00,0.00,0.20)$

$\mathbf{v_2}\_' = \mathbf{v_2}$

$S_0^{(4)} = 0.5498$

Run 10000

$\mathbf{v_1}' = (1.75,1.50,0.67,0.20,0.00,0.00,0.67,1.25,1.67)$

$\mathbf{v_2}' = \mathbf{v_2}$

$S_0^{(10000)} = 0.9017$

Nineteen out of 10000 runs resulted in a higher similarity between $\mathbf{v_1}'$ and $\mathbf{v_2}'$. The significance is therefore 0.0019, i.e. $\mathbf{v_1}$ and $\mathbf{v_2}$ are similar with a high probability.

Globularia punctata and *Anthyllis vulneraria*: a case study

Small populations of *Globularia puncata* and *Anthyllis vulneraria* have been screened over one vegetation period for natality and mortality of leaves (for more details on the data collection consult Hiebinger & Mucina, submitted). The results for the respective species are given in Fig. 1 (cumulative numbers of newly-born leaves in *Globularia punctata*) and Fig. 2 (cumulative numbers of dead leaves in *Anthyllis vulneraria*). A synchronity analysis (without considering a time shift) shows the leaf-natality patterns of *Globularia* (Table 1, upper matrix) to be highly synchronous, while the leaf-mortality patterns in *Anthyllis* (Table 1, lower matrix)

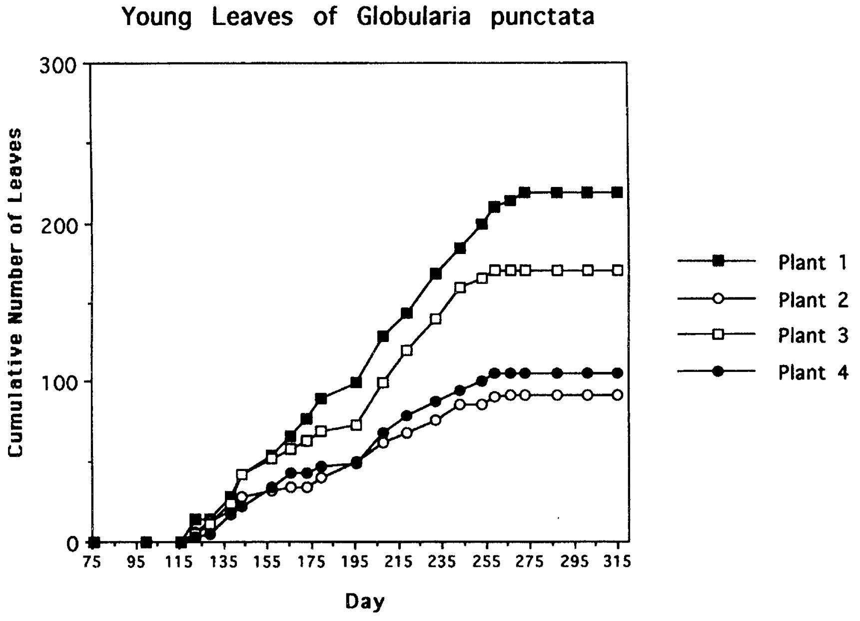

Figure 1. Cumulative number of young leaves of *Globularia punctata*.

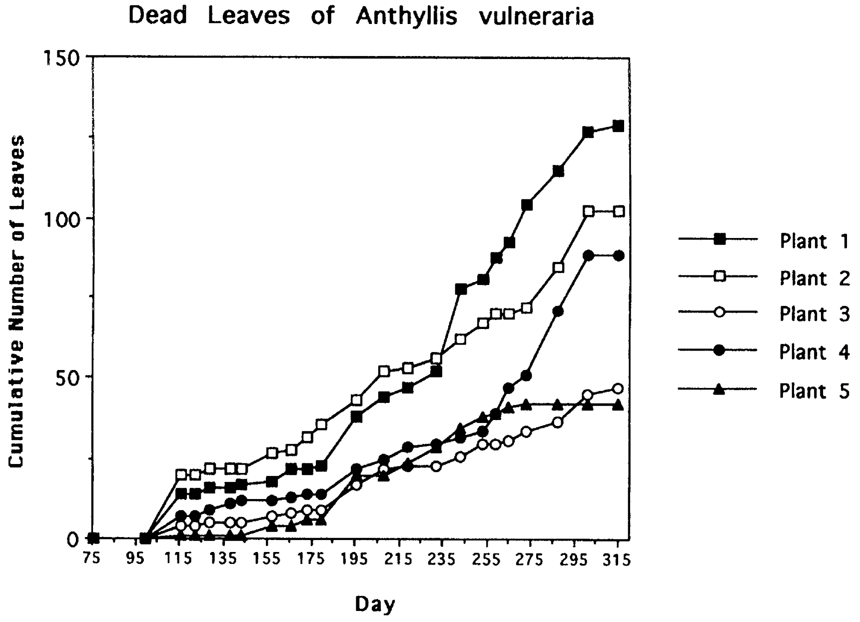

Figure 2. Cumulative number of dead leaves of *Anthyllis vulneraria*.

Table 1. Test for synchronity of the leaf-growth curves (both for young and dead leaves). The entries in the partial tables mean the probability of achieving a better correspondence between a reference curve and the compared one by means of Monte Carlo randomization procedure applied 1000 times to the compared curve. The bold-printed probabilities do not exceed the 5% confidence interval.

Globularia punctata (young leaves)

	1	2	3
1			
2	**0.006**		
3	**0.000**	**0.000**	
4	**0.000**	**0.006**	**0.000**

Anthyllis vulneraria (dead leaves)

	1	2	3	4
1				
2	0.081			
3	0.088	0.680		
4	0.090	0.120	0.300	
5	0.058	0.415	0.091	0.366

Table 2. Synchronity of respective pairs of individual (i,j) leaf-demography curves for the natality patterns in *Globularia punctata* and mortality patterns in *Anthyllis vulneraria*. No shift is indicated by (0), temporal shifts of the curve for species (j) backward/forward by 1, 2 and 3 sampling periods is indicated by negative and positive values, respectively. The highest resemblance values for a pair are given in bold.

Globularia punctata (young leaves)

i	j	-3	-2	-1	0	1	2	3
1	2	0.5852	0.5566	0.6777	**0.8321**	0.6738	0.8242	0.6649
1	3	0.6416	0.5765	0.6841	**0.8943**	0.7220	0.7442	0.6864
1	4	0.5874	0.7606	0.6714	**0.8908**	0.7555	0.7920	0.7236
2	3	0.5434	0.7596	0.6660	**0.8992**	0.5655	0.5349	0.3823
2	4	0.5158	0.8273	0.7484	**0.7953**	0.6743	0.5277	0.5222
3	4	0.4880	0.6945	0.7070	**0.8595**	0.7834	0.6120	0.5341

Anthyllis vulneraria (dead leaves)

i	j	-3	-2	-1	0	1	2	3
1	2	0.4108	**0.6766**	0.5858	0.6754	0.4484	0.5021	0.5769
1	3	0.4525	0.5441	**0.6746**	0.6629	0.5124	0.4457	0.6255
1	4	0.6813	**0.7434**	0.6772	0.6558	0.6181	0.5723	0.4093
1	5	0.4311	0.3069	0.4158	**0.6617**	0.5805	0.6205	0.4935
2	3	0.4591	0.5384	0.5624	0.7133	**0.7772**	0.4963	0.4245
2	4	0.4767	0.4854	0.5988	**0.6552**	0.6531	0.5516	0.6389
2	5	0.4628	0.5139	0.4098	0.4857	**0.5243**	0.5202	0.5045
3	4	0.3866	0.6280	0.6397	0.5762	**0.7263**	0.5711	0.5535
3	5	0.4783	0.4833	0.2816	**0.6250**	0.5812	0.5130	0.4587
4	5	0.2487	0.2926	0.2827	0.4571	0.3667	**0.6319**	0.5755

are asynchronous. The high synchronity between the curves of individual plants of *Globularia* was stressed also by a time-shift analysis (Table 2, upper matrix) showing the highest values of resemblance to occur when no time shift was simulated. With *Anthyllis* (Table 2, lower matrix), the time-shift yielded improvement in many compared pairs. Still, the values of resemblance failed to achieve even the lowest one found for *Globularia*. The high synchronity in leaf natality patterns in *Globularia* is in accordance with the high predictability of phenological patterns in this species (Hiebinger & Mucina, submitted). On the other hand, *Anthyllis* proved to be very unpredictable phenologically showing also rather haphazard patterns in leaf natality, flower and seed set (Sterk 1975, Sterk et al. 1982, Hiebinger & Mucina submitted).

Programme availability: The programmes (executable files; MS-DOS compatible computers) for the computation of synchronity of growth-curves is available cost-free (a non-formatted floppy should be sent first). Two sample data files are also supplied.

Acknowledgements. This work was assisted by a grant of the Austrian Science Foundation (FWF), Nr. P9072-BIO to L.M.

References

Hiebinger, C.K. & Mucina, L. submitted. Leaf demography patterns and reproduction timing of interstitial perennial herbs in a dry grassland.

Houle, G. 1994. Spatiotemporal patterns in the components of regeneration of four sympatric tree species - *Acer rubrum, A. saccharum, Betula alleghaniensis* and *Fagus grandifolia.* J.Ecol. 82: 39-53.

Mitchley, J. 1988. Control of relative abundance of perennials in chalk grassland in southern England. III. Shoot phenology. J. Ecol. 76: 607-616.

Sarukhàn, J. & Harper, J.L. 1973. Studies on plant demography: *Ranunculus repens* L., *R. bulbosus* L. and *R. acris* L.: I. Population flux and survivorship. J.Ecol. 61: 675-716.

Sterk, A.A. 1975. Demographic studies of *Anthyllis vulneraria* L. in the Netherlands. Acta Bot. Neerl. 24: 315-337

Sterk, A.A., van Duijkeren, A., Hogervorst, J. & Verbeek, E.D.M. 1982. Demographic studies of *Anthyllis vulneraria* L. in the Netherlands. II. Population density fluctuations and adaptation to arid conditions, seed populations, seedling mortality, and influences of the biocenosis demographic features. Acta Bot. Neerl. 31: 11-40.

Sydes, C.L. 1984. A comparative study of leaf demography in limestone grassland. J. Ecol. 72: 331-345.

van der Sman, A.J.M., Joosten, N.N. & Blom C.W.P.M. 1993. Flooding regimes and life-history characteristics of short-lived species in river forelands. J.Ecol. 81: 121-130.

Manuscript received: April 1995