COMMUNITY ECOLOGY 3(2): 125-146, 2002
1585-8553 © Akadémiai Kiadó, Budapest

Some interesting characteristics of the vegetation process

L. Orlóci[1,2], V. D. Pillar[1], M. Anand[3] and H. Behling[4]

[1] *Departamento de Ecologia, UFRGS, Porto Alegre RS, Brazil 91540-000. E-mail: vpillar@ecologia.ufrgs.br*
[2] *Department of Biology, the University of Western Ontario, London, Canada N6G 4V1*
E-mail: lorloci@uwo.ca
[3] *Department of Biology, Laurentian University, Sudbury, Canada P3E 2C6*
E-mail: manand@nickel.laurentian.ca
[4] *Center for Marine Tropical Ecology, Fahrenheitstrasse, 6, Bremen, Germany 28359*
E-mail: hbehling@zmt.uni-bremen.de

Keywords: Markov chains, Phase space, Trajectories.

Abstract: The characteristics discussed are measurable on the process "trajectory", the path traced out by vegetation transitions in time. It is argued that since the trajectory is symptomatic of factor influences and governing principles, it has to be an object of central interest in dynamic studies. Theoretical points, scaling scenarios, and analytical tools are the main topics. Numerical examples illustrate the applications.

Introduction

... this we do affirm —
that if truth is to be sought in
every division of Philosophy,
we must, before all else,
possess trustworthy principles
and methods for the discernment of truth.
— Sextus Empiricus

Our essay is devoted to the exploration of vegetation dynamics. Being one of possibly thousands already in circulation on the subject, the essay's novelty resides with the approach which emphasises the characteristics of the process trajectory. These include phase-structure, determinism and attractor migration, periodicity, dimensionality, and parallelism.

Phase structure implies the segmentation of a trajectory into significantly distinct intervals. Determinism, or directedness, is the manner in which the process could be running its course if the attractor, the set of conditions that define process direction, has not been overwhelmed by random effects. Periodicity is the tendency for certain structural variables, such as diversity for example, to revisit past states with regular or irregular frequency. The notion of dimensionality is understood in the sense of complexity, measured as a fractal, related to shape. Parallelism implies the tendency of the process to run its course in a manner of co-ordination with itself during different periods of time, or with trajectories from other sites.

Characteristics of this type are called *emergent* for the reason that their existence is tide to the community, not repeated at the level of any community component. Furthermore, being among the intangibles of the community, they cannot be measured directly, but rather, they have to be inferred analytically. We believe that emergent characteristics are sensitive indicators of factor influences, and as such, they are symptomatic of the governing principles of the vegetation process. Therefore, we believe that the in-depth study of emergent characteristics is not only fully warranted, but should prove to bring forth substantial progress toward the framing of an applicable general theory of vegetation development.

A three-step study scenario is historically invoked in the study of complex phenomena (Orlóci 2000). It seeks to answer the WHAT, HOW and WHY questions in that sequence (see Wiegert 1988, Anand 1994, 1997). Proceeding in the manner of such a study scenario is necessarily an exercise in *successive approximations*. It is that way because knowledge of complex phenomenon, being at its lowest and at its most intuitive at the beginning, will grow about itself by experience and as a consequence of step by step refinements in the scenario itself. These are hallmarks of successive approximation. Not surprisingly, the scenario is also hierarchical. The reason for this is that the understanding of complex phenomena progresses

through scales, historically from the macro-characteristics to the details. By moving down through scales, and not moving up, guarantees that the natural unity of the complex phenomenon, such as the vegetation process, is kept conceptually intact.

It is an interesting fact that Science always found utility of macro-manifestations in the search for causes and governing principles. To see this to be true, it is sufficient for one to think of the works of the Great Masters. Newton's basic laws of motion came about this way, so did the Darwin-Wallace theory of species evolution by natural selection[1], the Mendelian 3:1 statistical law of particle based inheritance and Kerner's doctrine[2] of plant community development by facilitation.

We propose to concentrate in this essay on the WHAT question. It is in that connection that we present details of a comprehensive methodology. Elaborated earlier (Orlóci and Orlóci 1988, Pillar and Orlóci 1996, Anand and Orlóci 1997, He and Orlóci 1998, Orlóci 2000, 2001a,b, Orlóci, Anand and Pillar 2002), the methodology is now expanded significantly through refinements that can only come from trials and collective thought. The reader's participation is, of course, invited.

Preliminary notions

We use the term "community level" when we describe the source of characters, the state of complexity, etc. Clearly, a scale or generalisation is implied, one or possibly several steps removed from the organismic and population functions. But we keep in mind that levels and functions are linked seamlessly into a system by natural laws and therefore their analytical isolation is arbitrary.

Isolation of levels and their analytical re-linking is an unalterable fact of practice. This is by necessity, since it is the only way to put a handle on measurability. In practice, this means the taking of characters from the lower levels, called *primitives*, and making them the axes of a reference space, called *phase space.* It is this space into which process trajectories are mapped by analytical projection, and examined for their characteristics, such as phase-structure, determinism and attractor migration, periodicity, complexity, and parallelism. Community level characteristics exist in other forms, of course, that can be measured directly. Radiation absorption and reflectivity are two examples.

Phase-structure, determinism and attractor migration, periodicity, complexity, and parallelism are at the apex of a character hierarchy. The following character arrangement is hierarchical:

— Process (phase-structure, determinism, attractor migration, periodicity, complexity, parallelism)
— Community (entropy, complexity, velocity)
— Population (biomass, photosynthetic type, facilitation)

As regards the analytical linking of the states of these characteristics directly to causes, we see it as in the first place a statistical problem. Of course, the statistical results may not help us beyond creating the framework within which specific deeper questions can be asked and further studies designed based on the knowledge gained. But this is quite in line with the practice of successive approximation. Naturally, we have to keep in mind that reasoning from the analytical results will be very much dependent on the time, space, and so-called analytical scales (Orlóci 2001a, b). Anchoring the scaling problem and the detection of general tendencies in specific examples are a characteristic of the paper.

The evidence

Among the many sources, data from *direct observation* of change over time and *paleopollen* and *tissue* records from sediments are most reliable. But, the most readily available sources are not these. They have their origin in phytosociological and other ecological surveys based on *space-for-time substitution* (Wildi and Schütz 2000).

It is intuitive that data from these sources will contain latent structures not obvious upon first inspection of a data set. To reveal these, one need access to appropriate analytical tools. We consider a "good" tool to be problem-specific. Yet, it should be unhindered by *a priori* unfulfilled assumptions. The latter may lead to the rejection of many well-known statistical techniques (Orlóci 1993, 2001b).

There are pitfalls in the use of observational data, and particularly, when they represent paleopollen and tissue counts. The main problem with these stems from the plant populations in the same sample not being identified at the same taxonomic level. This is by necessity, independent of choice. Other problems have to do with the scarcity of dated and calibrated samples (see Delcourt and Delcourt 1987, He and Orlóci 1998 and references therein).

Beyond the problems with a specific data source, it is randomness in the process that tends to muddle up the picture, and keep process characteristics obscured from di-

1 Explained in: Darwin, C. 1859. *The Origin of Species by Means of Natural Selection or the Preservation of Favoured Races in the Struggle for Life.* Murray, London.

2 Explained in his landmark book of 1863 "*Das Pflanzenleben der Donauländer*" published by Wagner in Innsbruck.

rect observation. The multi-scale nature of the process and multidimensionality add further to difficulties.

In our usage, the term *random* simply implies the superimposition of unruly, unpredictable behaviour upon a basic determinism that we expect to exist. *Stochasticity* is also implied in the context of the impossible exact *a priori* prediction of process states. In randomness thus lies one of the fundamental problems in the analytical probing of the vegetation process. We refer to Orlóci (2001a) and references therein for further thoughts on these.

Analyses in the manner of this essay traverse scales that are spatial and temporal. Furthermore, they are also related to the variable-list used and techniques applied. In these regards, the reader may be interested to consult materials presented elsewhere (Mandelbrot 1967, 1977, Schroeder 1991, Levin 1992, Schneider 1994, 2001, Orlóci 2000). "Multi-scale" implies that the manifestations of the process, or at least their perception, will depend on scale. "Multi-scale" does not need to imply, albeit it may be true that the process is naturally structured according to our choices of scale. In fact the process has natural unity and integrity, irrespective of its analytical divisibility. Yet, analytical division is useful because it helps to elucidate the characteristics of the process. What has been said underlies the importance of the combined use of reduction and synthesis in both analysis and reasoning.

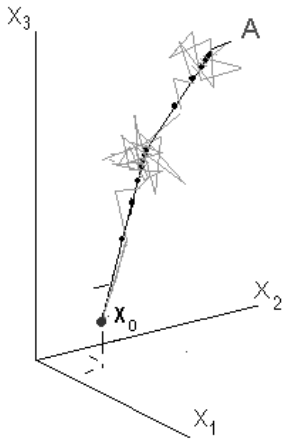


Figure 1. Phase space (X_1, X_2, X_3) with time axes ($\mathbf{X}_0$ to **A**). Point A represents the attractor conditions that define process direction in the present. Two fictitious processes are superimposed. One is the theoretical, moving Markov chain (dotted line, described in Appendix A) and the other (zigzag line) based directly on the hypothetical taxa. Symbol $\mathbf{X}_0$ implies the Null or Initial State, a point (X_{O1} X_{O2} X_{O3}) at which the observer's time begins.

Phase space

We deal with the process based on its mappings within *phase space.* The term "phase" emphasises the time related serial nature in the structures involved. We use records of taxa as co-ordinates.

The phase space need not have taxa for its dimensions. Any variable measured on organisms or populations could be used. At an extreme, such as in the example given by Bartha et al. (1998), Juhász-Nagy's information theoretical descriptors of coenostates represent the reference axes. Wildi (1998) introduced a further complexity into presentation of a reference space by making the trajectory points represent local vegetation maps. By applying Wildi's method, one can see how entire maps evolve in time. This is not unlike what is seen in Anand and Kadmon (2000) in connection with a study of Sinai vegetation.

A record set taken at a given time point is called an *observational vector* or *paleorelevé*. The use of discrete time points for recording, when the process itself is perfectly continuous, has the appearance of a serious anomaly. But there is no other practical way to record compositional changes in a community under field conditions over realistic periods of time.

It is practical to specify the phase space as a Cartesian space. It has the useful property of being linear with co-ordinates plotted on orthogonal axis. In such a space, non-linear structures appear as curved mappings. To identify the actual type of non-linearity that nature produces is important, because the type reflects back upon the behaviour of the process.

The alternative of making the axes curved to make non-linear structures appear linear, would not be practical, considering that success with this requires knowledge of the exact functional form of the non-linearity.

We present an idealised process trajectory in Figure 1[3]**.** The phase space has *four* axes. Three are taxa and the fourth is time. The basic number is usually much greater than three in practice. It must be kept in mind that every trajectory point is in the past, except point ***A****,* the attractor, that defines the direction of the process in the present. Of course, the attractor as a set of conditions in nature is in perpetual change. It is a known fact that the change is faster than the rate of readjustment in the process to equilibrate with it. Because of this, Nature can never attain the placid equilibrium that landscapes are seen to posses in renaissance painting (Groening 2001). Same goes for the early theorisations about the "vegetation climax".

Phase structure

Figures 2 and 3 present the mappings of realistic trajectories which we use to exemplify the property that we call *phase structure*. This remarkable property is seen in the way in which the trajectories are segmented into phases. A two-phase structure, linear and chaotic, is most clearly seen in Figure 2. The linear phase is suggestive of attractor stability over several time steps. The chaotic or non-linear phase indicates a rapidly migrating attractor.

3 We mention Feller (1957) as the premier reference on Markov chains. We also direct attention to Orlóci (2000) for review of related materials, and recommend the reading of Appendix A, which presents a semi-technical tour of essentials, before proceeding further with the text.

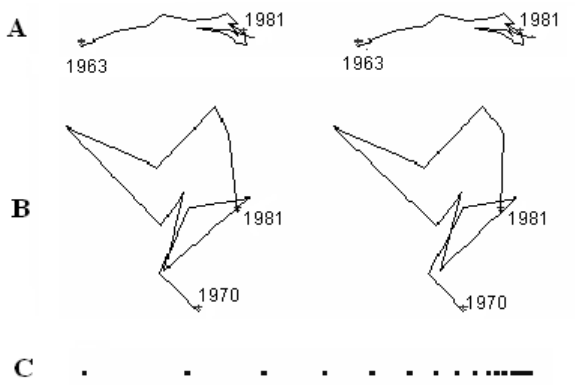


Figure 2. Eigenmapping of the Atlantic Heathland trajectory for stereo viewing. Graph A is based on the original data of De Smidt published in Lippe et al. (1985), reproduced in Table 1. The 2-dimensional mapping recovers 98% of variation in the original distance configuration. Graph B is an enlargement of the trajectory segment for the last 11 steps. Graph C is a mapping of the fitted Markov chain. Markov probabilities were determined by the method of Orlóci et al. (1993). See Appendix A and the main text for explanations.

Such an attractor manifests its effects in the manner of sharp turns and changing point densities on the trajectory.

What can be said about the cause of a sudden phase shift, such as around 1970 in Figure 2A? The original records in Table 1 reveal severe process occlusion occurring around that time. This appears in the manner of dramatic reduction of bare soil surface, signalling onset of a critical threshold in carrying capacity. Interestingly, bursts of linearity (more or less straight trajectory line segments) reoccur after occlusion, albeit only for short periods, punctuated each time by a sharp directional turn.

Phase changes are well-pronounced in the Cambará trajectory (Figure 3) as well. These appear to be climate induced over the long past up to about 1000 calibrated years BP (point 98, Figure 3) when anthropogenic effects also come into play, especially after about during the last 200 years (point 177, Figure 3).[4] Considering the excessive width of some time steps, corresponding to an otherwise perfectly uniform sampling interval (1 or 2 cm), the question of possible effect of the accuracy of interpolated dates on trajectory phasing has to be asked. It could be true that interpolation exaggerates the appearance of a phase structure. But then the apposite could also be true. The actual effect cannot be known, without accurate dat-

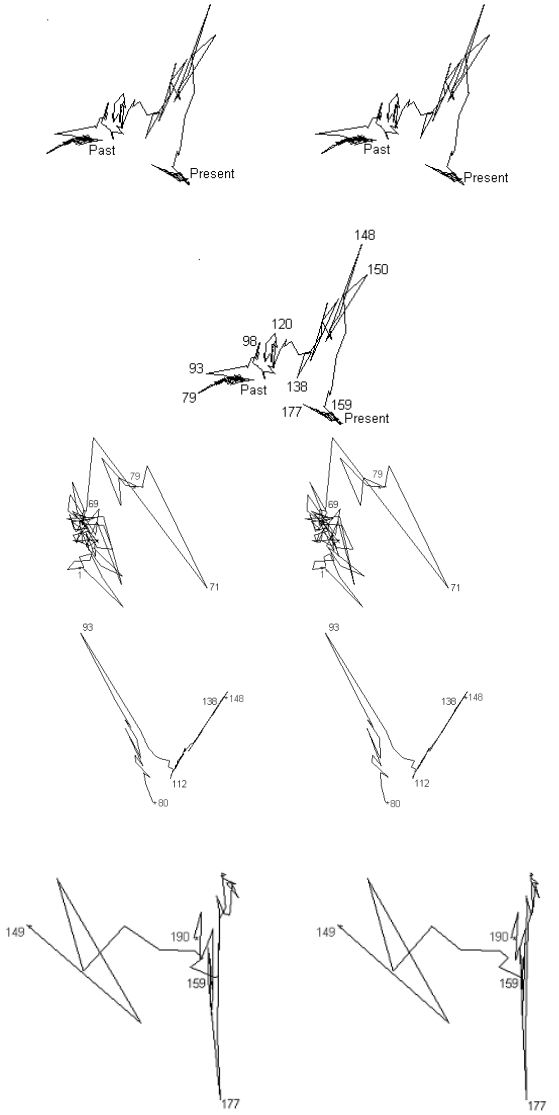


Figure 3. Eigenmapping of the Cambará trajectory for stereo viewing. First set of three graphs: 3-dimensional map captures over 89% of variation in the original distance configuration. The data set is H. Behling's (Behling et al. 2002). It contains high-precision paleopollen percentages for 11 pollen types in 190 sediment samples taken at each 1 cm from 0 to 80 cm depth, and at each 2 cm from 80 cm to the bottom. Dates are interpolations based on 7 AMS radiocarbon dates. The point marked "Past" corresponds to the bottom sediment horizon laid down over 42,000 ^{14}C years BP (before 1950). Bottom three pairs of graphs: Each pair represents a separate Eigenmap. Corresponding dates to numbers: 1 - 42840; 69 - 27690; 71 - 26640; 79 - 22470; 80 - 21950; 93 - 12570; 98 - 8890 14C years BP (10250 calibrated years BP); 112 - 3690; 120 - 2770; 138 - 975; 148 - 732; 149 - 708; 150 - 683; 159 - 465; 177 - 159; 190 - -50 (2000 AD). From point 98 to the present all dates are in calibrated years BP.

4 See Behling et al. (2002) for more details on the evolution of the anthropogenic effect.

Table 1. Percentage point cover for bare ground and plant populations in 19 annual surveys of Atlantic Heathland vegetation undergoing recovery from the effects of severe fire and grazing. Data set from Lippe et al. (1985). Code: BG - bare ground, EN - *Empetrum nigrum*, CV - *Calluna vulgaris*, ET - *Erica tetralix*, MC - *Molinia caerulea*, CP - *Carex pilulifera*, JS - *Juncus squarosus*, RA - *Rumex acetosella*, OS - other species. Note reversals in the process in 1977 and 1978.

Year	BG	EN	CV	ET	MC	CP	JS	RA	OS
1963	57.1	17.9	8.6	11.6	0.0	0.2	0.0	4.7	0.0
1964	44.0	25.0	13.7	12.2	0.0	1.1	0.2	3.9	0.0
1965	32.7	34.9	13.9	14.3	0.0	0.5	0.0	3.7	0.0
1966	27.5	36.8	20.0	14.1	0.1	0.9	0.2	0.3	0.1
1967	19.7	46.1	21.0	10.8	0.1	0.7	0.4	0.5	0.7
1968	10.7	54.2	22.2	10.6	0.7	0.6	0.4	0.0	0.5
1969	6.7	55.7	23.3	10.4	0.3	2.0	0.7	0.1	0.7
1970	5.8	61.1	23.7	6.9	0.2	1.2	0.7	0.2	0.3
1971	9.5	57.6	24.7	6.6	0.4	0.6	0.4	0.0	0.3
1972	8.4	62.1	23.7	3.6	0.3	1.2	0.1	0.0	0.6
1973	4.4	67.9	21.3	3.3	0.2	0.6	0.4	0.0	2.0
1974	8.5	58.1	25.8	4.7	0.6	1.3	0.7	0.0	0.4
1975	9.2	62.2	24.3	2.5	0.6	0.9	0.2	0.0	0.1
1976	9.9	58.2	24.9	3.7	0.6	1.1	0.7	0.0	1.0
1977	19.6	48.4	23.5	5.7	0.3	1.2	0.4	0.1	0.9
1978	12.1	58.1	22.7	4.8	0.4	0.4	0.0	0.2	1.3
1979	9.3	65.1	20.3	2.7	0.0	1.5	0.1	0.2	0.9
1980	7.3	68.2	21.5	1.2	0.5	1.0	0.1	0.1	0.2
1981	5.4	65.5	20.8	4.6	1.0	1.6	0.4	0.3	0.6

ing performed on many more core samples. It should, however, be clear that low-intensity dating is not a theory based choice, but a practical (mainly monetary) necessity.

The question whether we are dealing with a pure Brownian trajectory in which trends arose, as they can do in the nature of a chaotic process or we are seeing a stochastic process playing itself out in which random variation from all sources is superimposed on a basic deterministic trend, is legitimate. In other words, it is of theoretical importance to contemplate if we are observing chaos at work or dealing with a special case of stochastics (Orlóci 2001a). We feel that we are seeing the manifestations of stochastics. We see support for this in the appearance of a strong, underlying deterministic component of compositional changes upon which random oscillations are imposed. As regards the Cambará trajectory, the long-term determinism appears in greater part induced by non-random, non-feedback type climate shifts that tend to move the global conditions between extremes through long periods of gradual, albeit frequently disrupted, cooling or similar warming events, not the least as a mechanistic consequence of cyclic type celestial mechanics described in connection with the Ice Age Phenomenon (Milankovitch 1941).

So far we did not address the *modus operandi* in the delimitation of trajectory phases. Edge detection is involved, in our examples mainly based on gestalt, but in others on edge detection algorithms. I this respect we refer to the literature (see Orlóci and Orlóci 1990 and references therein).

Determinism

The techniques to be outlined help the question to be answered:

"How smoothly has a trajectory run its course within a selected time period, in the sense of steadfastly keeping its time-forward momentum".

If there is much chaotic directional change obvious in the trajectory, the assumption of a simple determinism will be undermined. Three things are to be expected in the long run:

- Ecological and other biological laws make it impossible for the process to be either totally deterministic or totally chaotic.
- As a consequence of this, stochastics is likely to prevail over the long-run. That is to say, random variation will not obliterate for long the basic deterministic trend.
- Could the case of the fractal not prevail? We see evidence of a level of chaoticity in phases. But even then, order could arise by chance (Anand and Orlóci 1997).

Different techniques are available to quantify the level of determinism. A common feature in these is reference to a "standard", a presumed type of determinism that would prevail if some stipulated conditions existed in nature. We describe cases in the following table:

Standard	Criterion	Decision
Markov chain	Stress of two distance configurations, one Markov and the other based on paleorelevés. The Markov chain is fitted to the records, but then it is also fitted to the randomised records to generate a probability distribution for stress under the chance compositional transitions hypothesis. (See Appendix A)	Given an observed stress value $\sigma(\boldsymbol{D}_{observed};\boldsymbol{D}_{Markov})$ where $\boldsymbol{D}_{observed}$ represents a distance matrix based on the original records $\boldsymbol{X}$ and $\boldsymbol{D}_{Markov}$ a distance matrix based on Markov scores $\boldsymbol{PX}_t$, accept the Markov type determinism hypothesis if $\sigma(\boldsymbol{D}_{observed};\boldsymbol{D}_{Markov})$ is less or equal to σ_α which is a chosen probability point of the distribution of $\sigma(\Delta_{random};\Delta_{Markov})$ under the random transitions hypothesis. In the latter, Δ_{random} represents a distance matrix based on the randomised original records and Δ_{Markov} a distance matrix based on Markov scores fitted to the randomised records. The probability point σ_α is such that α proportion of the values in the distribution of $\sigma(\Delta_{random};\Delta_{Markov})$ is equal to or less then σ_α. In other words, small values of sigma in probability terms discriminate in favour of Markov type determinism. Note that α is a probability in the left tail of the distribution.
Time ordinals	Rank correlation of two distance configurations, one based on time ordinals and the other on the observed records.	Correlation has zero expectation. Large correlation values discriminate in favour of determinism.
Time axis	Topological similarity of two trajectories, one representing the time axis and the other the observed trajectory. See Appendix B.	Large similarity values discriminate in favour of a simple time-related determinism. Note, the value of the mean under random transitions is the topological similarity coefficient's expected value.

The tests of significance incorporate elements of Goodall's (1964) reasoning about probabilistic similarity measures, but the actual implementations are different from his. We use randomisation (Monte Carlo) experiments rather than a theoretical distribution to determine probabilities.[5] Randomisation experiments invoke the proposition that the vegetation process is driven by random transitions, which is the antithesis of determinism.

Markov trajectory a standard

We assume at this point that the reader is familiar with the contents of Appendix A. Figure 2C presents an example of a Markov trajectory. Obviously, it is completely deterministic, albeit the progression through it is not linear, which is seen from the crowding of states as time passes, as long as the transition probability matrix $\boldsymbol{P}$ stays constant. Having said this much, the reader will no doubt be lead to the conclusion that the strength of directedness in the natural process can be measured by using the fitted Markov trajectory as standard (see Orlóci et al. 1993, Orlóci 2000). We have done this and present some results in Table 2.

Clearly, Table 2 indicates that Markovity is not a far-fledged assumption in the linear phase, but untenable in the chaotic phase. This fact warns not to generalise about the Markov chain without prior investigation of the process trajectory's phase structure.

Set of ordinals a standard

The ordinals 1, 2, 3, … are used to indicate a simple chronological progression. It is unlikely that determinism will be found in nature. But the point is to have a fixed configuration for comparison. Taking the rank correlation coefficient, symbolically $r(\mathbf{D}_c;\mathbf{D}_t)$, a value is computed between the elements of two matrices. One of the matrices is $\mathbf{D}_c$, containing the compositional distances of paleorelevé pairs, and the other $\mathbf{D}_t$, containing the paleorelevé pairs' chronological differences.

Sample correlation values are given in Table 3. Inspection reveals an almost functional relationship of time and compositional change in the linear phase (1-8) of the Atlantic Heathland trajectory, and a remarkable lack of it in the turbulent phase (1970 to 1981). The same conclusion can be drawn in the case of Cambará. The expectations are consistent with the random (RND) case, gener-

5 See fundamentals and applications in Edgington (1987, Pillar and Orlóci 1996).

Table 2. Levels of Markov type determinism in the process trajectories as identified. See the text and Appendices A for description of the method. Probabilities were determined in randomisation testing. Periods in the second column are chosen to show the method's capabilities.

	Period length	Observed stress $\sigma(D_{observed};D_{Markov})$	Expected stress under H_0	Variance of the expectation	Discrimination in favour of H_0, *i.e.* the process not being Markov type* α	Decision
Atlantic Heathland (Lippe et al. 1985)	1963-1981	0.2435	1.0685	$1.^{4}1896$**	<0.001	Markovity upheld
Atlantic Heathland (Lippe et al. 1985)	1963-1970	0.1618	0.8094	$1.^{3}2603$	<0.001	Markovity upheld
Atlantic Heathland (Lippe et al. 1985)	1970-1981	1.4508	1.4279	$1.^{5}9013$	0.635	Markovity rejected
Cambará from step 189 to 214	26	0.4423	0.9935	$0.^{4}3774$	<0.01	Markovity upheld
Cambará from step 1– to 214	214***	1.6426	1.6211	$0.^{5}2905$	>0.94	Markovity rejected

* Compositional transitions Chance-ruled. α is a value on the 0 to 1 interval.
** Exponent following decimal point gives the number of zeros as in 1.00001896.
*** Trajectory transformation to equal time steps (200 years each) applied. This transformation is needed in Markov analysis.

Table 3. Trajectory determinism, measured as a rank correlation of time based and composition based distances as explained in the text. Symbol RND designates an artificial trajectory consistent with random compositional transitions. Expectations and probabilities are determined in randomisation experiments based on 1000 RND trajectories. The probability of an at least as large rank correlation value as the observed is given.

Site	Observed rank correlation	Expected correlation under H_0	Variance of the expectation	P
Atlantic Heathland 19 yrs	0.5810	0.0005	0.0612	<0.001
Atlantic Heathland, first 8 yrs	0.9291	0.0007	0.0377	<0.001
Atlantic Heathland, last 12 yrs	0.2938	-0.0069	0.0142	0.537
Cambará 42,000 yrs	0.4787	0.0069	0.0027	<0.010
RND	-0.004	0.000	0.006	0.920

ated by allowing compositional transitions to be driven by chance.

Time axis a standard

We assume the familiarity with the contents of Appendix B. The method to be used is the same, except for one of the trajectories, which is the time axis. The question of how parallel is the given process trajectory to the time axis is to be answered.

The topological similarity graphs and the deviation graphs are presented in Figure 4. A convex shape of curve "a" above the confidence limits (b,d) is the characteristic to be watched. Some comments on this are in order:

- The topological similarity coefficient (curve "a") is a direct expression of a trajectory's forward parallelism with another trajectory, in this case the time axis.
- Any portion of the topological similarity graph above the upper confidence limit (curve "b") implies that the similarity has reached significance at least as high or higher than the 95% exclusion. That is to say, determinism appears significantly more intense than could be expected should chance completely ruled the compositional transitions. The strength of the similarity is directly proportional to the deviation from expectation (curve "c"). Any portion of the topological similarity graph below the lower confi-

Table 4. Characteristic points of the topological similarity coefficient (TC curves "a", Figure 4) and their description. See details in the main text and in Appendix B.

Trajectory	Tolerance limit %	Observed value *TC*	Expected value of *TC* under random compositional transitions	Deviation from expectation	Variance of the expectation	Lower 0.95 confidence limit	Upper 0.95 confidence limit
Hanging Lake	7	.5105	.446	.0644	.0006	.3987	.4933
	50	.9382	.3861	.552	.0029	.2813	.4909
Cambará	7	.4849	.446	.0388	.0006	.3987	.4933
	44	.9189	.36	.5588	.0028	.2572	.4628
Lagoa das Patas	13	.4632	.4044	.0587	.0006	.3539	.4549
	50	.8779	.3861	.4918	.0029	.2813	.4909

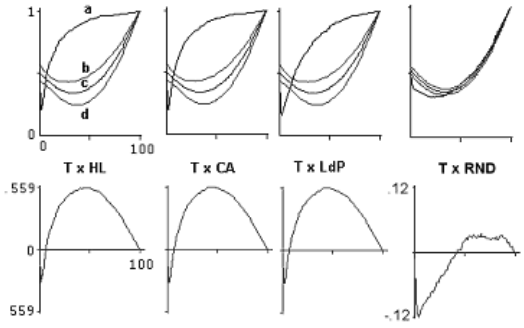


Figure 4. Topological similarity of the Hanging Lake (HL), Cambará (CA), Lagoa das Patas (LdP), and random (RND) trajectory vis-a-vis the time axis. Top row: vertical scale, topological similarity; horizontal scale, tolerance limit increasing from 0 to 100% in 1% steps. Bottom row: deviations from random expectation. Symbols: a - topological similarity as measured (see Appendix B); b - upper limit of 0.95 probability confidence interval; c - random expectation; d - lower limit of 0.95 probability confidence interval. Limits and expectations are determined in randomisation experiments. The graphs use the 15-taxon subsets from the HL and LdP data sets, and the 11-types subset from the CA data set (see Appendix D). Note that portions of graph "a" above the upper confidence limit "b" indicate concordance higher than random expectation. Portions below the lower confidence belt d indicates weaker concordance than random expectation. Further, the dropping of graph "a" below "d" in the TxRND comparison is appears to be an artefact, owing to an imperfectly random RND**.** See Appendix D, for address to HL and LdP data sources.

dence limit (curve "d") implies that the similarity is significantly lower than random expectation.

- Three points have special significance on graph "a". One is the point where it rises above the confidence belt. This happens to be at relatively low values of the tolerance limit in the examples. The other is the point where it crosses the lower confidence limit. The third is where it deviates maximally from random expectation ("c"). Some results are given in Table 4. We note the TxRND curve, which establishes the expected behaviour of the topological coefficient. The other graphs are very similar. This gives substantive support to the von Post (1946) theorem (see section on Paralellism and Appendix B).

We observe the concave graphs (b, c, d) in Figure 4, and put a first question "Why this orderliness?" It is relevant that the concave graphs came about in the averaging of large numbers of *TC* values applied to random trajectories that we generated by random permutation of the observed co-ordinate data. Should this orderliness surprise us? Not really, considering that we have in hands the case of the "fractal" (Orlóci 1991). In other words, natural laws dictate order to arise in chaos if some orderly perturbation is allowed to undermine its integrity. The catalyst in our case is the step by step broadening of the tolerance limits. This causes the value of *TC* to start deviating from 0.5, the theoretical value at 0 tolerance limit. It is quite apparent that the chaotic structure collapses as the tolerance limits broaden and aggregates are formed, which will first lower the values of *TC,* implying that the trajectories starts appearing less and less similar, before the change reverses direction and the values of *TC* begin to rise. The question is "Why this way, why not on a convex path?" The "easy" answer is regard this a governing principle, as it were, inherent in the collapse of a chaotic structure, one which becomes revealed by experiments with large collections and large numbers of iterations (narrow increments of the tolerance radius). These thought should challenge the thoughtful student whom we expect to communicate back to us.

Periodicity

It is clear from what has already been presented that syndynamics behaves in manners that can be described as periodicity, the return of community characteristics to previously attained states. A variable that in time varies in the manner of a series of sinus waves is an example of "periodicity". But it is clear that in the nature that we are addressing periodic change is always irregular for both wavelength and amplitude. A simple example of periodic variation is the turnover of taxa in time. In good part, however, complex structures are involved whose state vari-

ables are one or more steps removed from taxonomic composition and whose dynamics is asynchronous with population turnover. Typical state variables in this category have to do with trajectory behaviour, such as velocity, diversity, and complexity (Orlóci 1991, 1998, 2000, Anand and Orlóci 1996, 2000, He and Orlóci 1998, Orlóci et al. 2002).

Velocity is defined as the amount of change in community composition per unit time. This can be calculated in the manner of in which *d(j,j+1)* is the compositional distance over which the process moved within the time interval from t_j to t_{j+1}. The velocity values are trajectory specific and for that reason they are comparable only as ranks between cases.

The specific family of entropy functions of interest is discussed elsewhere in general (Orlóci1991 and references therein) and in a companion paper (Orlóci et al. 2002) from which we take specific cases. The reader should note that we use the term "diversity" in a rather technical sense, either as entropy of order one (disorder-based diversity) or entropy of order zero (the upper limit for disorder-based diversity, also known as "richness"). We use Anand's complexity measure (Anand and Orlóci 1997, 2000, Anand 2000, Orlóci et al. 2002), specifically the structural component of total complexity or Kolmogorov's algorithmic complexity, which is coding theory based and as such the complexity measured is proportional to the level of "difficulty" in a community's description. A useful bit of information to know about is that the upper limit of Kolmogorov's complexity is maximum entropy which occurs when the frequency distribution is an equidistribution.

There are many ways to define complexity. Design (Dawkins 1986), organisation (Fosberg 1965), shape (Mandelbrot 1967, 1977), and coding related complexity (Anand 2000, Anand and Orlóci 1996, 2000, Orlóci 2000) are éclat examples. We find particularly useful for our purposes Kolmogorov's definition (see Anand and Orlóci 1996, 2000). This makes the measured level of complexity conditional on the code length $L=\Delta+H_1$, or in other words, the level of the difficulty in an object's description. Since H_1 is disorder related, the Δ term captures that portion of complexity that is not disorder but structure related. As expected, Δ is zero when a distribution on which it is defined is most dispersed, and maximal when the distribution is least dispersed. Other measures of complexity may target the *shape* of objects in other than coding theoretical term. This is so when a frequency distribution's *skewness* and *kurtosis* is measured, or a graph's *fractal dimension* is determined in the manner of Mandelbrot (1967, 1977).

Figures 5 and 6 give examples from two sites. Interesting questions about these address the relationship of periodic change intrinsic in velocity, entropy and complexity, and extrinsic to these in the environment. We plan to focus on the intrinsic part and leave the extrinsic to a companion paper (Behling et al. 2002).

Regarding the patterns of change, the ECV graphs in Figures 5 and 6 indicate some interesting tendencies. In

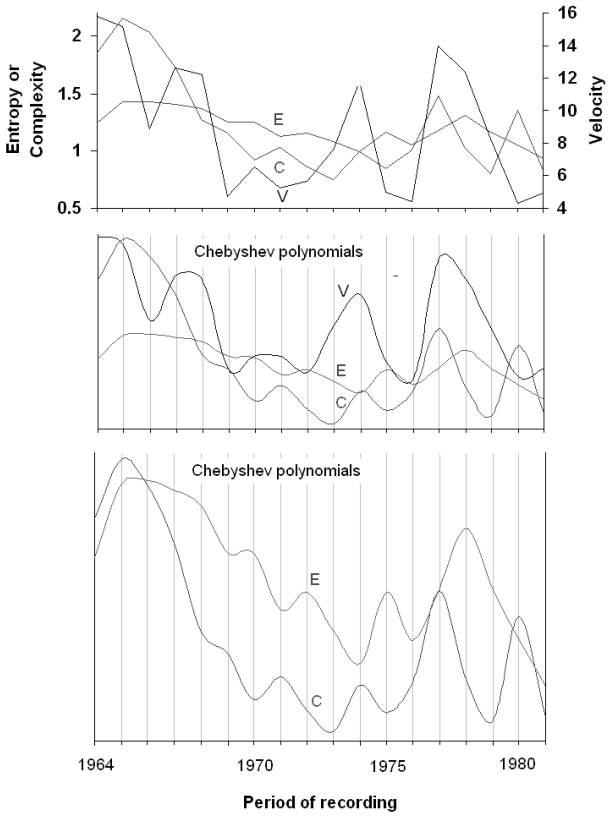


Figure 5. Composition based diversity (entropy of order one, E), structural complexity (C) and process velocity (V) in relation to time in the Atlantic Heathland site (see Lippe at al. 1985 and our Table 1). The lower two graphs show smoothing by fitted Chebyshev polynomials (order 16). Events to be noted: 1964 - emergence from log-term heavy grazing and fire; 1970 - severe reduction of bare ground, occlusion reached; 1976 - short draught period. Note: vertical scales are arbitrary and for that reason the graphs are comparable only by their topology (topography) and not by their magnitudes.

the Atlantic Heathland entropy rises after fire and intensive grazing around 1963, reaching local maximum soon after. From about 1966, entropy undergoes a steady decline, reaching a minimum around 1974. At around 1971, severe reduction of the open ground surface in the site begins to show its effects (see data in Table 1). This causes a shift from the linear initial phase to a non-linear phase. With this shift, a concomitant rise of entropy dominates the trend until 1978. Velocity appears to be accelerated by periodic perturbations, followed by sharp deceleration.

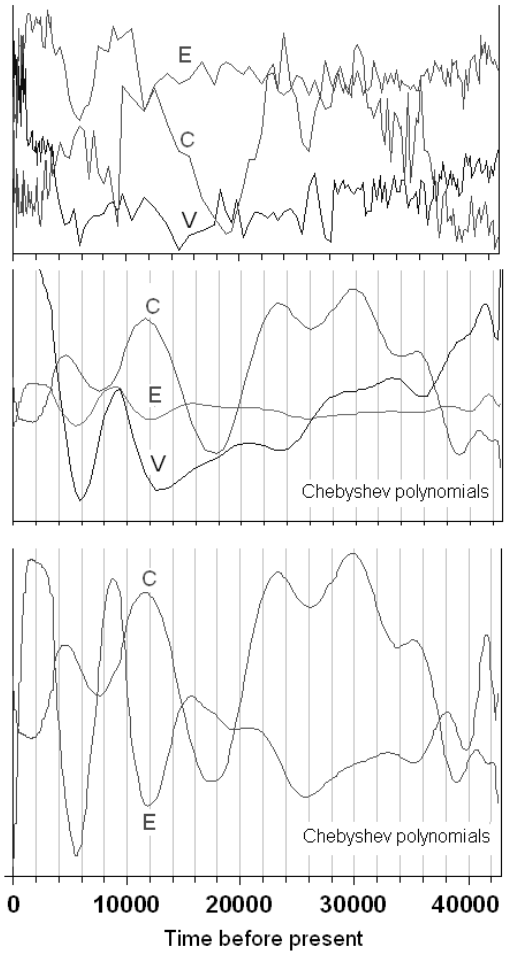


Figure 6. Compositional diversity (entropy of order one, E), structural complexity (C) in relation to process velocity (compositional change in unit time, V) in the Cambará site over the time period as indicated. Graphs on top are based on Behling's 11pollen types over 190 sampling depths. The lower two graphs are Chebyshev polynomials of order 20 fitted to original graphs. Velocity is plotted on a logarithmic scale. Vertical scales are adjusted to enhance the graphs readability. Some significant time point approximation based on inspection of temperature graphs (see Webb 1992) and so-called bio-indicators (after Behling et al. 2002): 39,000 ^{14}C years BP - intense climate cooling begins; 23,000 BP - coldest part of last Ice Age approached; 18,000 BP - Ice Age ending; 12,000 BP - cooling cycle ending; 6,000 BP - Hypsothermal; 2000 BP -Little Climatic Optimum; 280 BP - deep point in Little Ice Age reached.

Table 5. Values of the correlation coefficient for the trajectories in Figure 5 after Chebyshev smoothing within periods as indicated.

Window width as a time interval	r(V,E)	r(V,C)	r(E,C)
1964-1970	0.3477	0.7034	0.6352
1970-1978	-0.052	0.66761	0.2616
1978-1981	0.9355	-0.2701	0.0034
1964-1981	0.5364	0.6351	0.6758

Legend: V – velocity; E – entropy of order one; C – structural complexity

Structural complexity seems to follow a similar pattern as entropy, but with sharp reversals after 1971.

The patterns in the Cambará ECV graphs appear to be dominated by the Late-Quaternary global temperature fluctuations (see graphs in Webb 1992), and structural complexity appears to be a premier indicator of these. Clearly, structural complexity increases at times of climate cooling and decreases during climate warming. Process velocity shows an opposite trend: deceleration during climate cooling and acceleration during climate warming. Similar tendencies are seen in entropy as well. Superimposed on the main trends are oscillations owing to environmental perturbation, possibly the effects of major volcanic activity on the global climate and of anthropogenic activity on the vegetation of the region in the more recent past (Behling et al. 2002).

An interesting aspect of the graphs in Figures 5 and 6 is their co-ordination. We probed the graphs for this by use of a correlation analysis based on the product moment correlation coefficient. The choice is purely a matter of convenience. Other correlation measures could be used and possibly other aspects of the relationships would be uncovered. The method required the laying of contiguous windows of given length on the graphs within which the correlation coefficient is computed. The windows were sited regularly or at random, and so was also the choice of window lengths. Taking a few selected cases under the option of regular siting and non-random lengths, in the Atlantic Heathland case we have for the Chebyshev polynomials (Figure 5) the values in Table 5. Note that smoothing by Chebyshev fitting removes much chaotic variation, and the correlations are expected to be affected.

What do the graphs and correlations tell us about periodicity in the Atlantic Heathland? Obviously, the change is irregular in time and somewhat chaotic. The overall correlation is reasonable high. We find the large-number distribution of correlation for the Chebyshev polynomials interesting. These are given in Table 6. An important thing is the diagnostic value of correlations for perturbation. Table 7 contains the correlations for this. Obviously, the correlations are symptomatic of the effects

Table 6. Large-number distribution of the correlation coefficient for the Chebyshev polynomials (Figure 5) generated by randomly positioned windows of random length.

Case *,*	r(*,*)		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0	Total	Zero
V,E	>0	f	668	906	1706	1788	673	435	1043	487	161	0	7867	169
		%	6.6	9	17	17.8	6.7	4.3	10.4	4.8	1.6	0	78.6	1.6
	<0	f	409	152	87	0	0	402	335	0	320	259	1964	
		%	4	1.5	.8	0	0	4	3.3	0	3.2	2.5	19.6	
V,C	>=0	f	406	540	1438	745	1380	2898	1272	147	0	0	8826	169
		%	4	5.4	14.3	7.4	13.8	28.9	12.7	1.4	0	0	88.2	1.6
	<0	f	518	87	234	0	0	85	81	0	0	0	1005	
		%	5.1	.8	2.3	0	0	.8	.8	0	0	0	10	
E,C	>=0	f	1393	868	587	229	353	1223	1381	1493	254	0	7781	871
		%	13.9	8.6	5.8	2.2	3.5	12.2	13.8	14.9	2.5	0	77.8	8.7
	<0	f	421	177	341	85	159	165	0	0	0	0	1348	
		%	4.2	1.7	3.4	.8	1.5	1.6	0	0	0	0	13.4	

Note: Windows are randomly laid. V – velocity; E – entropy of order one; C – structural complexity, f - frequency.

Table 7. Values of the correlation coefficient of the trajectories in Figure 5 after Chebyshev smoothing within the periods as indicated.

Period	Precursor to period	r(V;E)	r(V;C)	r(E;C)
1964-66	Severe fire and grazing	-0.6033	-0.1579	0.8829
1966-69	Calluna expansion	0.98563	0.6997	0.8103
1970-75	Severe reduction of bare ground	-0.8239	0.1258	-0.1230
1976-81	Severe draught	0.8037	0.3594	0.1566

V – velocity; E – entropy of order one; C – structural complexity

Table 8. Values of the correlation coefficient for the trajectories in Figure 6 after Chebyshev smoothing within periods as indicated.

Window width by ^{14}C years BP	r(V,E)	r(V,C)	r(E,C)
42841- -50	0.0144	-0.4439	-0.3971
42841-39000	0.4752	-0.1542	0.2395
39000-30000	0.831	-0.8265	-0.903
30000-23000	0.1476	0.4779	0.8615
23000-17000	-0.5411	0.555	-0.7998
17000-12000	0.8257	-0.9846	-0.9082
12000-7000	0.9415	-0.5114	-0.7503
7000-4000	0.6261	0.4573	-0.401
4000-2000	0.9617	-0.9757	-0.9123
2000-500	-0.8733	0.9229	-0.9648
500- -50	-0.6432	0.3084	-0.9193

Legend: V – velocity; E – entropy of order one; C – structural complexity

of disturbance, that trigger abundance changes in *Calluna,* reduction of bare ground, etc.

The product moment correlation values for the Chebyshev polynomials for Cambará (Figure 6) are given for selected periods in Table 8 and the large-number distribution of the correlation coefficient in Table 9. Clearly, the correlation is very much period specific and in many cases remarkably high. In this case, as in the previous one, the reader may speculate about the biological significance of the distributions whether the numbers have predictive value. But regardless of this, the main utility of correlation analysis is the precise fixing of structures as they ex-

Table 9. Large-number distribution of the correlation coefficient for the Chebyshev polynomials (Figure 6) generated by randomly positioned windows of random length.

Case *,*	r(*,*)		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0	Total	Zero
V,E	>0	f	495	487	452	522	612	795	1531	1141	835	0	6870	560
		%	4.9	4.8	4.5	5.2	6.1	7.9	15.3	11.4	8.3	0	68.7	5.6
	<0	f	288	218	193	134	102	105	119	145	572	694	2570	0
		%	2.8	2.1	1.9	1.3	1	1	1.1	1.4	5.7	6.9	25.7	0
V,C	>=0	f	193	177	218	335	278	215	262	233	450	0	2361	299
		%	1.9	1.7	2.1	3.3	2.7	2.1	2.6	2.3	4.5	0	23.6	2.9
	<0	f	272	300	414	489	1029	1710	1201	501	806	618	7340	0
		%	2.7	3	4.1	4.8	10.2	17.1	12	5	8	6.1	73.4	0
E,C	>=0	f	73	47	61	51	21	21	23	27	86	0	410	81
		%	.7	.4	.6	.5	.2	.2	.2	.2	.8	0	4.1	.8
	<0	f	101	118	229	650	1804	1695	1363	1398	556	159	9509	0
		%	1	1.1	2.2	6.5	18	16.9	13.6	13.9	5.5	15.9	95	0

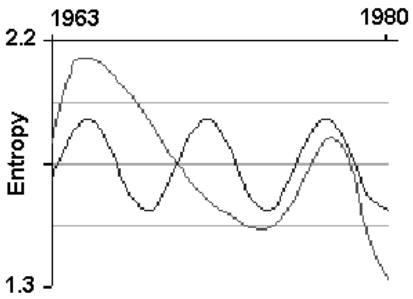
V – velocity; H – entropy of order one; C – structural complexity

Table 10. Numerical results from the comparison of process trajectories in Figure 7. See technical details in Appendix B and the main text.

Trajectory pair	Length of common trajectory segment in 200 yr steps	Tolerance limit % at maximum deviation	Maximum deviation	Observed topological similarity at maximum deviation	Expected value	Lower 95% confidence limit	Upper 95% confidence limit	Decision on parallelism
Hanging Lake x Cambará	206	50	0.513	0.8991	0.3861	0.2813	0.4909	Significant
Hanging Lake x Lagoa das Patas	206	50	0.4398	0.8260	0.3861	0.2813	0.4909	Significant
Cambará x Lagoa das Patas	214	56	0.4504	0.8701	0.4196	0.3167	0.5225	Significant

isted. The magnitude of future change can only be gauged by reference back to these structures.

We should note that the graphs that we use to illustrate periodicity could be subjected to direct probing for the precise numerical nature of periodicity. An entire literature treats the problem area under various titles, such as spectral analysis, time series analysis, autocorrelation analysis, and so on. We refer for example and references to Fox and Parker (1968), Rayner (1971), He and Orlóci (1998), and Orlóci et al. (2002) from whom we reproduce the following graph:



The irregular curve is a Fourier polynomial of order 16. At this order, the fit is extremely precise. But is it the type on which periodic properties should be measured? It could be based on such characteristic values as wave frequency, amplitude, phase angle, phase shift, and so forth. But there is a big problem with polynomials, namely, they cannot be used reliably outside the data range. To see this, it is sufficient to consider the case of the same type of polynomial fitted to the same entropy graph shortened just by one time step as Orlóci et al. (2002) have done. The predicted entropy for 1980 becomes 2.14 bits. This is a point far exceeding the 1980 level in the original curve (1.34 bits), and far above the observed value (1.55 bits). The sinusoid curve shown is the second residual obtained in the third step of the decomposition of the original entropy graph. The decomposition is such that the sum of all stepwise sinusoid curves recreates exactly the observed entropy graph. The decomposition statistics (up to the 5th residual) are as follows:

Residual #	0	1	2	3	4	5
Wave frequency:	1	2	3	4	5	1
Wave amplitude:	.19	.15	.13	.08	.04	.04
% of total entropy accounted for	45	24	18	6	2	2
Cumulative %:	45	69	87	93	95	97

Table 11. Length-related fractal dimension of the VEC graphs. Refer to the text and Appendix C for technique and discussions. The fractal dimension computed has theoretical limits between 1 and 2. Adjustment of the graphs to 100 maxima on each axes were applied. Table contains maxima as explained in Appendix C. Basic data are given in Table 1 for Atlantic Heathland.

Graph	Velocity		Entropy		Structural complexity	
	D	*MD*	*D*	*MD*	*D*	*MD*
Cambará	1.764	1.441	1.567	1.555	1.747	1.606
Cambará (Chebyshev)	1.173	1.098	1.291	1.244	1.341	1.255
Atlantic Heathland	1.356	1.239	1.368	1.229	1.320	1.296
Atl. Heathland (Chebyshev)	1.352	1.247	1.367	1.228	1.321	1.297

D -length-related fractal dimension, *MD* – running mean fractal dimension..

The numbers indicate rapid decline in wave amplitude and in the amount of entropy in the residuals. Our choice of the 2nd residual serves a purely cosmetic purpose, namely the assurance to have a reasonable number of waves covered. In any case, the sinusoid curve is too regular and as such it should not be expected to be a reliable descriptor of the natural diversity process. It should be clear that the characteristic values of a sinusoid curve will change depending on the period length and the residual examined. It is to be noted also that while the idea of "best fit" is operational with sinusoid curves, it is absolutely not so with polynomials. There can be a unique wave frequency at which the sinusoid fit to the observations is best, but there is no uniqueness in the same sense with polynomials whose precision of fit is expected to increase by simply increasing the polynomial's order.

Parallelism

We now turn to the von Post (1946) hypothesis. Based on this, using appropriate techniques, we should be able to detect a global parallel tendency in vegetation development, irrespective of locality but respective of time. The technical details of our method and relevant references are discussed in Appendix B. Results are presented for four trajectories in Figure 7. We give some numerical results in Table 10.

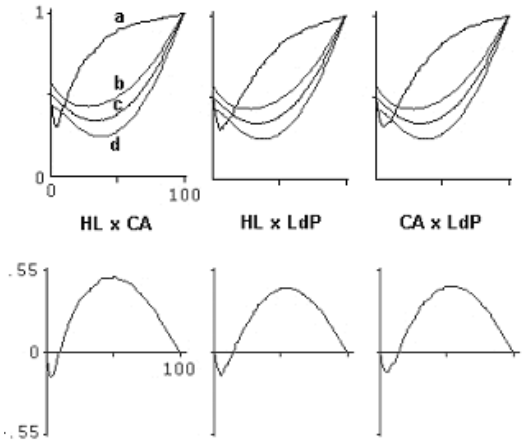


Figure 7. Graphs of the topological similarity coefficient in top row, and deviations from random expectation in bottom row. Vertical axis: similarity or deviation. Horizontal axis: tolerance limit %. Transformation to equal time steps in the trajectories (200 years) is applied. Symbols: a - topological similarity graph; b - upper limit of 95% confidence interval; c - random expectation; d - lower limit of 95% confidence interval; HL - Hanging Lake, Yukon Territory; CA - Cambará, Rio Grande do Sul; LdP - Lagoa das Patas, Amazonas.

Interpretation of the graphs follows the points discussed already in connection with Figure 4. In those terms, significant co-ordination is found in all three cases. Taking into account all the above, we conclude that the von Post hypothesis of process parallelism is in fact true.

Dimensionality

The dimension of a graph, any of those in Figures 5 and 6, can be taken to mean different things. It can be the number of reference axes (2 in the example), the number that describes the magnitude of the values plotted (velocity, entropy, complexity), or something entirely different, and not even expressed as an integer number, but a *fractional*.

A specific case should help to make some interesting points about graphs and the complexity of their shape. The source reference for theory is Mandelbrot (1967, 1977, Schroeder 1991). We use a graph length related fractal dimension, which we outlined in Appendix C. Related applications and references are described elsewhere (Orlóci 2001b, Orlóci et al. 2002).

We begin with results given in Table 11 for graphs in Figures 5 and 6. The graphs were adjusted to equal size (100 maximum on both axes). The initial calliper width was set equal to 1 and incremented in unit steps up to 100. The conclusion emerging from the numbers is a ranking of the graphs according to their shape complexity. Fractal dimension was determined before and after smoothing by high-order Chebyshev polynomials. Since the consequence of smoothing is the removal of random oscillations and a reduction of complexity, it is better to focus on the graphs as they come without smoothing.

Some remarks are in order. The length-related fractal dimension is a measure of shape complexity. This is quite clear from the reduction of the fractal dimension in the Cambará graph following Chebyshev smoothing. Smoothing did not have the same effect in the Atlantic Heathland graph. This was expected, since the original graph itself is not complex. Note that a straight line has fractal dimension 1, and a graph created by random oscillations over a long period has fractal dimension two (see Appendix C). Cambará graphs are measurably more complex than the Lippe graphs, but their high complexity is related to excessive random variation.

Concluding remarks

The point has been made that the community process has its own level of functioning. This was not meant to say that the vegetation process can be understood without understanding the lower level processes. But understanding starts with the precise description of the process on its own level. In line with this thinking, our work promotes a methodology by way of which the community process can be probed for regularities. The methodology targets phase structure, determinism, periodicity, parallelism, and dimensionality. There are several good reasons that mediate the choice of these:

- They are symptomatic of causes and governing principles.
- They are applicable over any range of floras or environmental conditions.
- When precisely described, their effect on the thinking about the process in problem terms is synergetic.

As far as the approach is concerned, it certainly facilitates the structured analysis of the vegetation process in the manner of successive approximation. We do not need to place limitations on what is essentially a probing and learning process by incorporation of specific mechanisms or primitives *a priori*. Vast quantities of data can be laid open, available from past surveys and paleoecological studies that could not be analysed otherwise to their full information potential in process trajectory studies.

Acknowledgements: Juliana Cunha and Carolina Blanco assisted in the early stage of method development. We thank them for their effort. We express our special gratitude to Sándor Bartha for the in-depth review that lead to substantial improvements of clarity. L. O. expresses his gratitude to UFRGS and CNPq for hospitality and generous support.

References

Anand, M. 1994. Pattern, process and mechanism. The fundamentals of scientific inquiry applied to vegetation science. *Coenoses* 9: 81-92.

Anand, M. 1997. Vegetation dynamics: toward a unifying theory. Ph.D. thesis. Department of Plant Sciences, The University of Western Ontario, London, Canada.

Anand, M. 2000. The fundamentals of vegetation change: complexity rules. *Acta Biotheoretica* 48: 1-14.

Anand, M. and R. Kadmon. 2000. Community-level analysis spatiotemporal plant dynamics. *Ecoscience* 7:101-110.

Anand, M. and L. Orlóci. 1996. Complexity in plant communities: notion and quantification. *J. theor. Biol.* 101: 529-540.

Anand, M. and L. Orlóci. 1997. Chaotic dynamics in multispecies community. *Env. and Ecol. Stat.* 4: 337-344.

Anand, M., and L.Orlóci. 2000. On partitioning of an ecological complexity function. *Ecol. Model.* 132:51-62.

Bartha, S., T. Czárán, and J. Podani. 1998. Exploring plant community dynamics in abstract coenostate space. *Abstracta Botanica* 22:49-66.

Behling, H., V. D. Pillar, L. Orlóci and S.G. Bauermannd. 2002. Late Quaternary *Araucaria* forest, grassland (Campos), fire and climate dynamics, inferred from a high resolution pollen record of Cambará do Sul in southern Brazil. (Manuscript.)

Dawkins, R. 1986. *The Blind Watchmaker.* Penguin Books, Suffolk.

Delcourt, P. A. and H. R. Delcourt. 1987. *Long-term Forest Dynamics of the Temperate Zone.* Ecological Studies 63, Springer–Verlag, New York.

Edgington, E.S. 1987. *Randomization Tests.* 2nd ed. Marcel Dekker, New York.

Feller, W. 1957. *An Introduction to Probability Theory and its Applications*. Vol. I, Wiley and Sons, London.

Fosberg, F. R. 1965. The entropy concept in ecology. In: *Symposium on Ecological Research in Humid Tropics Vegetation*, UNESCO and Government of Sarawak, Kuching, Sarawak. pp. 157-163.

Fox, L. and I. B. Parker. 1968. *Chebyshev Polynomials in Numerical Analysis.* Oxford University Press, London.

Goodall, D.W. 1964. A probabilistic similarity index. *Nature* 203:1098.

Groening, G. 2001. About landscape, nature and culture. Lecture Presented at the 2001Congress of the Brazilian Ecological Society in Porto Alegre, Brazil.

He, X.S. and L. Orlóci. 1998. Anderson Pond revisited: the late Quaternary vegetation process. *Abstracta Botanica* 22:81-93.

Levin, S.A. 1992. The problem of pattern and scale in ecology. *Ecology* 73: 1943-1967.

Lippe, E., De Smidt, J.T and D.C. Glen-Lewin. 1985. Markov models and succession: a test from a heathland in the Netherlands. *J. Ecol.* 73: 775-791.

Lorenz, E.N. 1963. Deterministic nonperiodic flow. *J. Atm. Sci.* 20: 130-141.

Mandelbrot, B. B. 1967. How long is the coast line of Britain? Statistical self similarity and fractional dimension. *Science* 156: 636-638.

Mandelbrot. B.B. 1977. *Fractals: Form, Chance and Dimension.* Freeman, San Francisco.

Milankovitch, M.M. 1941. Canon of isolation and the Ice-Age problem. Royal Serb Academy Special Publication 133.

Orlóci, L. 1991. *Entropy and Information.* Ecological Computations Series: Vol. 3. SPB Academic Publishing, The Hague.

Orlóci, L. 1993. The complexities and scenarios of ecosystem analysis. In: G. P. Patil and C. R. Rao (eds.), *Multivariate Environ-*

mental Statistics, North Holland/Elsevier, New York. pp. 421-430.

Orlóci, L. 1998. Appendix In: G. Fekete, K. Virágh, R. Aszalós and L. Orlóci. 1998. Landscape and ecological differentiation of *Brachypodium pinnatum* grasslands in Hungary. *Coenoses* 13:39-53.

Orlóci, L. 2000. From Order to Causes. A personal view, concerning the principles of syndynamics. Published at the web address: http://mywebsite.netscape.com/lorloci/koa

Orlóci, L. 2001a. Pattern dynamics: an essay concerning principles, techniques, and applications. *Community Ecol.* 2: 1-15.

Orlóci, L. 2001b. Perspects and expectations: reflections on a science in change. *Community Ecol.* 2: 187-196.

Orlóci L., Anand, M. and X.S. He. 1993. Markov chain: a realistic model for temporal coenosere? *Biometrie-Praximetrie* 33: 7-26.

Orlóci, L., Anand, M. and V. D. Pillar. 2002. Biodiversity analysis: issues, concepts, techniques. Community Ecol. 3: 217-236.

Orlóci, L. and M. Orlóci. 1988. On recovery, Markov chains and canonical analysis. *Ecology* 69: 1260-1265.

Orlóci, L. and M. Orlóci. 1990. Edge detection in vegetation: Jornada revisited. *J. Veg. Sci.* 1:311-324.

Pillar, V. D. and L. Orlóci. 1996. On randomisation testing in vegetation science: multifactor comparison of relevé groups. *J. Veg. Sci.* 7: 587-592.

Post, von L. 1946. The prospects for pollen analysis in the study of the Earth climatic history. *New Phytol.* 45:193-217.

Rayner, J. N. 1971. *An Introduction to Spectral Analysis.* Pion, London.

Schneider, D.C. 1994. *Quantitative Ecology: Spatial and Temporal Scaling.* Academic Press, San Diego. pp. 709-727.

Schneider, D.C. 2001. Concept and effect of scale. In: S. A. Levin (ed.), *Encyclopedia of Biodiversity.* Academic Press, San Diego.

Schönemann, P. H. and Carroll, R. M. 1970. Fitting one matrix to another under choice of a central dilation and a rigid motion. *Psychometrika* 35: 245-256.

Schroeder, M. 1991. *Fractals, Chaos, Power Laws.* Freeman, New York.

Selling, O.H. 1948. *Studies in Hawaiian Pollen Statistics. Part III. On the Late Quaternary History of the Hawaiian Vegetation*. Bernice P. Bishop Museum, Honolulu, Hawaii.

Webb, T. 1992. Past changes in vegetation and climate: lessons for the future. In: R.L. Peters and T.E. Lovejoy (eds.), *Global Warming and Biological Diversity*. Yale University Press, New Haven, London. pp. 59-75.

Wiegert, R.G. 1988. Holism and reductionism in ecology: hypothesis, scales and system models. *Oikos* 53:267-273.

Wildi, O. 1998. Simulating vegetation at a local scale. *Abstracta Botanica* 22: 3-11.

Wildi, O. and M. Schütz. 2000. Reconstruction of a long-term recovery process from pasture to forest. *Community Ecol.* 1: 25-32.

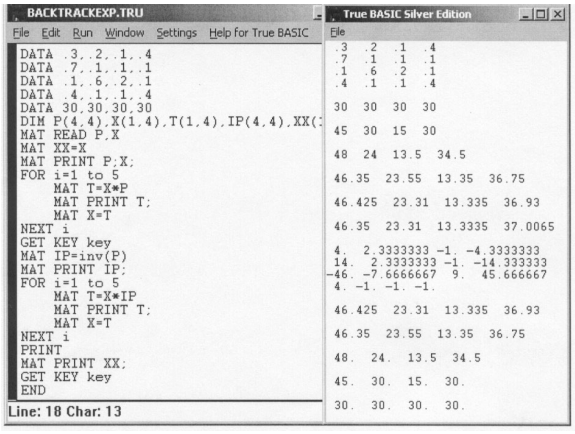
Appendix A

Any series of numbers (trajectory) that behaves according to the simple recursive relationship $\boldsymbol{X}_{t+1}= \boldsymbol{X}_t\boldsymbol{P}$ is said to be a Markov chain. In our case, $\boldsymbol{X}_t$ is an *s*-valued state vector at a point in time *t.* Since it is a state from the vegetations past, we refer to $\boldsymbol{X}_t$ as a paleorelevé. Symbol *s* can be thought of as the number of taxa that make up the relevé. $\boldsymbol{P}$ designates an *sxs* transition probability matrix. This matrix is such that the entry in the intersection of row *h* and column *i* is the probability of a given taxon *h* (row entity) being replaced by itself *(h=i)* or by another taxon $(h \neq i)$ in the next time step. Accordingly, the sum of the probabilities in any row of $\boldsymbol{P}$ is exactly one. The sum of the elements in any column need not be that way.

To reflect further on the Markov chain, we consider a case where a single taxon is involved. So $\boldsymbol{X}_o$ and $\boldsymbol{P}$ are both one-valued. Given $\boldsymbol{X}_0$ = [10] and $\boldsymbol{P}$ = [0.5], the Markov chain unfolds like this: 10 5 2.5 … Carrying on in the same manner, we have $\boldsymbol{X}_{10}$=0.01953125. The series converges to a value after a while that will not change significantly, no mater how many new steps are calculated. In practice, the value will depend on the precision of the instrument that performs the calculations.

The basic problem in Markov chain applications is determination of the transition probabilities. Transition probabilities cannot be foretold with any accuracy under the usual ecological conditions. They have to be obtained from data as postdictions. Orlóci et al. (1993) describe a practical solution to the problem (the first as far as we know) in the context of survey–type, time-series vegetation data.

It should be remembered that the stationary Markov chain (see Feller 1957), a chain with constant $\boldsymbol{P}$, requires two things to be completely defined: $\boldsymbol{X}_o$ and $\boldsymbol{P}$. As regards the frequently asked question whether or not the Markov chain has memory, the answer is definitely yes for path. This is easily seen from the existence of backtracking, namely $\boldsymbol{X}_{t-1}= \boldsymbol{X}_t\boldsymbol{P}^{-1}$. To see this in numbers, study the following algorithm and numerical results, and then contemplate the constraints under which backtracking is possible:



The code is on the left side and the output on the right. The top block of numbers in the output contains the con-

stant transition probabilities (***P***). The initial $\boldsymbol{X}_0$ is the vector [30 30 30 30]. Iterations produce the next set of numbers in 5 rows. The inverse of ***P*** is the central block of 4x4 numbers. The numbers in the last 5 rows shows that $\boldsymbol{X}_o$ is recoverable by post–multiplications of $\boldsymbol{X}_t$ by $\boldsymbol{P}^{-1}$. But under what conditions could this be done? —

a) There is no return from the "stable state" once it is reached. Of course the perception of stability depends on to the accuracy of the computer. It is deemed to have been reached when further iterations will leave $\boldsymbol{X}_t$ unchanged, so that $\boldsymbol{X}_t = \boldsymbol{X}_{t+1} = \boldsymbol{X}_{t+2} \ldots$.

b) Backtracking beyond the initial state $\boldsymbol{X}_o$ into a fictitious past is not biologically defined. Computationally, it produce numerical results revealing of some unique properties of the function irrelevant to the process' past.

c) Exact backtracking is an obvious impossibility in Nature. This is not surprising, considering that natural carbon copies of the vegetation are never produced.

d) The Markov chain, as we apply it to vegetation data, has discrete states, but the process itself is continuous. This is not an anomaly that the novice to the field should be tempted to remedy. The use of discrete states is a necessity, a consequence of the way in which the vegetation process can be observed in the field: in discrete steps.

e) The transition probabilities are under random and other types of effects in Nature. For this reason, a constant ***P***, or equivalently, the stationary Markov chain is an impossible state of existence in natural communities.

f) Since there is no facility built into the recursive equation to control the step size, equal time steps have to be assumed. In practice, equal time steps are created by interpolation if time is measured as a continuous variable. The method of interpolation used in this paper is as follows:

(i) Take a set of observations like those given here:

Interval in cm	0-10	10-20	20-30	Total
Interval in time units	5	3	7	15
Pollen count for taxon A	15	6	42	63
Average count	3	2	6	—

(ii) Arrange the averages into a 15-valued string like this,

3 3 3 3 3 2 2 2 6 6 6 6 6 6 6

(iii) If the time interval is set at, say 3, then obtain the interpolated pollen counts in the manner of the following:

Interval in time units	3	3	3	3	3	15
Pollen count for taxon A	9	8	10	18	18	63

Clearly, interpolation leads to smoothing, and in that sense a loss of information, but as a payoff, to possibly a clarification of trends.

g) The Markov chain converges to a point attractor consistent with ***P*** but independent of $\boldsymbol{X}_o$.

h) Any state $\boldsymbol{X}_t$ of a Markov chain is somewhat more distant from the previous state $\boldsymbol{X}_{t-1}$ than it is from the next state forward $\boldsymbol{X}_{t+1}$ (see Figure 2C).

i) When $\boldsymbol{X}_t$ is a paleorelevé, the series $\boldsymbol{X}_0, \boldsymbol{X}_1, \boldsymbol{X}_2, \ldots$ defines the sample process trajectory for which ***A*** is a target (attractor). Such an idea of the trajectory invokes notions from ballistics. The community's state is the projectile, propelled through time in the direction of a target ***A***. The conditions that set bounds on ***A*** constitute the "attractor" conditions. The attractor is a fixed target for the stationary Markov chain, and for such things as a perfect pendulum, but a moving target is more realistic in vegetation dynamics.

j) The phase space mapping of a Markov trajectory is like the entire graph Figure 2C. The position-by-velocity graph of the pendulum is a decaying spiral. The trajectory of community dynamics is like in Figures 2A or 3. The Markov chain in graph Figure 2C is fitted to graph A in the same figure. The Markov null state in the example is De Smidt's 1963 relevé, $\boldsymbol{X}_O$ = (57.1 17.9 8.6 11.6 0.0 0.2 0.0 4.7 0.0). The transition probabilities, determined by the method of gains and losses (see Orlóci et al.1993), are the elements in the matrix:

.76480	.13945	.05628	.02656	.00116	.00455	.00148	.00339	.00234
.01323	.95375	.01660	.00688	.00154	.00344	.00143	.00045	.00268
.01304	.05180	.92028	.00606	.00134	.00325	.00130	.00044	.00250
.01924	.11058	.04133	.81922	.00128	.00413	.00146	.00060	.00217
.02685	.12406	.04602	.01709	.77525	.00476	.00188	.00063	.00345
.03093	.13603	.05083	.01703	.00158	.75754	.00172	.00149	.00284
.04221	.15201	.06296	.01743	.00191	.00460	.71411	.00139	.00338
.09055	.17091	.07742	.04942	.00067	.00549	.00139	.60338	.00077
.02908	.18118	.06217	.01409	.00186	.00509	.00235	.00054	.70364

k) The attractor in reality oscillates under random influences or may even move into new regions of phase space under general change of conditions. When this happens ***P*** is changed. Points in time at which dramatic changes occur in ***P*** are causing sharp turns in the trajectory. In the trajectory segments where

Table 12. Co-ordinate data (a) and directional scores (b) corresponding to the trajectories drawn in Figure 8. Legend: pa - past; pr - present.

A or B				C				D				E				F			
a	b	a	b	a	b	a	b	a	b	a	b	a	b	a	b	a	b	a	b
1pa		1		8		8		4		3						13		14	
2	+	2	+	7	-	7	-	8	+	5	+					25	+	21	+
3	+	3	+	6	-	6	-	2	-	4	-	12		26		21	-	07	-
4	+	4	+	5	-	5	-	6	+	2	-	38	+	45	+	48	+	20	+
5	+	5	+	4	-	4	-	3	-	7	+	31	-	28	-	42	-	11	-
6	+	6	+	3	-	3	-	1	-	1	-	59	+	27	-	68	+	10	-
7	+	7	+	2	-	2	-	5	+	8	+	74	+	23	-	90	+	09	-
8pr	+	8	+	1	-	1	-	7	+	6	-	91	+	26	+	120	+	10	+

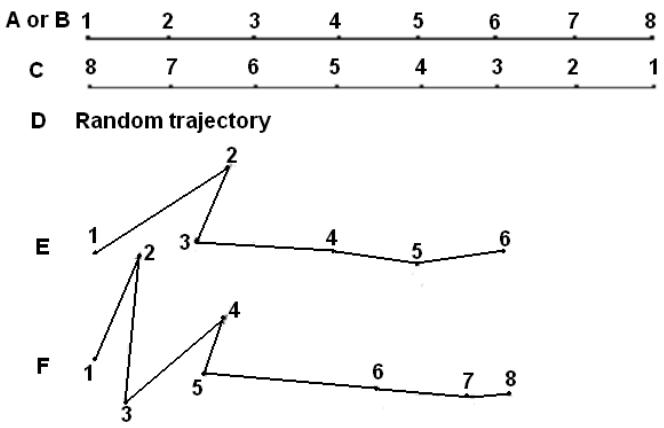


Figure 8. Sample trajectories corresponding to the data in Table 12. The case of time ordered forward is repeated twice (A,B), backward (C). A set of randomly defined points represent the "random" trajectory (D), not drawn. Trajectories E and F correspond to non-random co-ordinates. The example is adopted from Orlóci (1998) with revisions.

this keeps happening continuously, the Markov chain becomes a "moving Markov chain".

l) When the elements of ***P*** are under random effects, the attractor becomes some unpredictable point. Such an attractor is called a "strange attractor" (Lorenz 1963). It is noted that the cumulative effect of small random perturbations will blur both the appearance of Markov type state crowding on the trajectory as well as process directedness. Clearly, process predictability is hindered by random oscillations in ***P***. But unpredictability cannot be made total. This is because of basic ecological laws that make completely random compositional transitions an impossible state of vegetation.

Appendix B

Von Post (1944) examined some composite paleopollen diagrams from different sites and discovered a reoccurring regularity in the diagrams' morphology. This gave him the idea of parallel vegetation development in different regions. Selling (1948) may be consulted for interesting applications of the von Post method in the Hawaiian Islands. The von Post approach is an ingenious way to compare pollen diagrams, but it is not the most convenient one (on first site at least) for probabilistic testing of regional parallelism. The testing is more easily done in direct comparisons of trajectory mappings. The method to be described measures parallelism as the level of co-ordination determined based on the point-by-point similarity of trajectories in the manner used by Orlóci (1998, 2000, 2001a,b, He and Orlóci 1998).

We constructed artificial trajectories in Figure 8 on which we illustrate the methods application. These include time forward (A,B) and backwards (C). Each trajectory in the example has two sets of co–ordinates, which we give in Table 12**.** The comparison scalar for any trajectory pair is a topological similarity coefficient[6] $TC = M/p(s\text{-}1)$. The value of TC ranges from 0 to 1. The theoretical expectation of TC is taken to be 0.5 under a flat-chance law, all states from 0 to 1 having an equal chance of occurring) under random compositional transitions. Values indicate concordance above 0.5 and discordance

6 "Topological" is used with the explicit purpose to link the notion of similarity with the general idea of topography. "Topology" is then the study of the topography of objects as in land surveys, landscape ecology, trajectory morphometrics, etc. "Topology" is definitely not intended here to imply the mathematician's geometry on a "rubber sheet" as S. Smale has used the term to make vivid the nature of the mathematical concept.

below 0.5. Other symbols: M - number of matching directional scores (+, –, 0) on the axes; p - number of axes (taxa) of the reference space for the trajectory with the lesser number; s - elapsed period in equal time steps for the trajectory with the lesser period length.

As the computations progress, trajectory co-ordinates are transformed into forward (+) or backward (-) scores, or left to remain 0, if no change occurred. Following this, matches are counted. The scoring follows specific rules:

a) "Present" refers to the same time point or the same narrow time interval in all trajectories. When working with palaeopollen samples, the top sediment horizon may miss being from the "present", owing to possible perturbation by acts of erosion, decomposition, or other processes active in the solum. Because of this, it makes sense to repeat the comparison of the trajectories at varying lag of shifts forward or backward. The aim is to maximise *TC*.

b) In the interest of finding comparable points on different trajectory lines, we assume uniform time step width.

c) Noting the above, any given state of the process, symbolically **X**, is endowed with the potent property of being the outcome of random effects, and as such, of being at a locus of phase space unpredictable before the fact. It is justified, therefore, to regard any **X** as a member of the point swarm, any member of which could have materialised if chance dictated it to be that way. One cannot know exactly the radius of the swarm, but one can make provision for it by expanding the tolerance sphere[7] in a flexible manner around **X**. This is done starting with radius 0, incremented in small steps up to some arbitrary upper percentage limit of the co–ordinate values. All co-ordinates on all axes undergo tolerance transformations with the following possible results:

"0" is assigned if the tolerance belt around the given point includes the other compared point. If it does not, then:
"+" is assigned if the co-ordinate value of the given point is greater than the co-ordinate value of the chronologically next (forward) point on the dual axes. The meaning of the dual axis is explained in d) below.
"-" is assigned if the above relation is reversed.

d) When matches are counted, reference axes have to be paired between the trajectories. We refer to the members of a selected pair as "dual axes". There is no natural way for pairing the axes of different reference spaces. To overcome this problem, finding dual pairs may be in a manner that will maximise the value of *TC*. A method is described in the example below, which pre-empts the so called Procrustes manipulations of Shönemann and Carroll (1970).

e) Trajectories are compared in pairs. If one member of the pair is a time trajectory, the selection of duals is not a problem. Otherwise axes for pairing are purposefully selected to maximise *TC*. In the present example, the choices are based on local maxima within the 2 x 2 submatrices of between–axes correlations as in the following

Axis	B1	B2	C1	C2	D1	D2	E1	E2	F1	F2
A1	1.00	1.00	-1.00	-1.00	0.00	0.38	0.94	-0.43	0.95	-0.55
A2	1.00	1.00	-1.00	-1.00	0.00	0.38	0.94	-0.43	0.95	-0.55
B1			-1.00	-1.00	0.00	0.38	0.94	-0.43	0.95	-0.55
B2			-1.00	-1.00	0.00	0.38	0.94	-0.43	0.95	-0.55
C1					0.00	-.038	-0.94	0.43	-0.95	0.55
C2					0.00	-.038	-0.94	0.43	-0.95	0.55
D1							0.59	-0.09	0.21	0.55
D2							0.23	-0.52	0.29	-0.22
E1									0.99	-0.07
E2									-0.27	0.89

A quick example will clarify the arithmetic in the calculation of *TC*. Consider trajectories E and F. The M quantity of the shared directional scores is equal to 10. Since $p = 2$ and $s = 6$, $TC = 1.0$ or 100%. This implies that the two trajectories are indistinguishable in terms of their topology. The entire set of similarity values and probabilities (α) for the trajectories portrayed above are as follows:

Trajectory pair:	*AB*	*AC*	*AD*	*AE*	*AF*	Expectation		Bias in iteration
						Theoretical	iterated	
TC	1.00	0.00	0.50	0.60	0.57			
α	0.003	1.00	0.63	0.33	0.44			
	BC	*BD*	*BE*	*BF*	*CD*			
TC	0.00	0.50	0.60	0.57	0.50			
α	1.00	0.63	0.33	0.45	0.63			
	CE	*CF*	*DE*	*DF*	*EF*			
TC	0.4	0.43	0.60	0.64	0.70	0.500	0.516	+0.016
α	0.81	0.74	0.33	0.27	0.176			

7 "Tolerance sphere" should not be confused with the probabilistic "confidence sphere, belt, or interval" as used in statistical estimation. While confidence spheres or belts help to handle the sampling error, tolerance spheres, belts, or intervals help incorporate the possible effects of process stochasticity when making comparisons between the trajectory points.

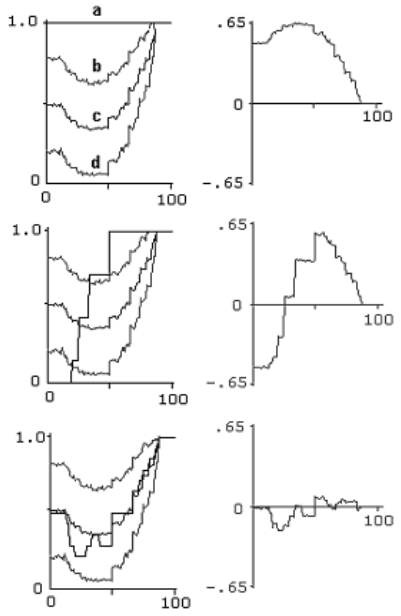


Figure 9. Graphs of the topological similarity coefficient (*TC*) in left column: a -observed value; c - random expectation (iterated), b,d - limits of the 0.95 confidence belt. Graphs of deviations (a-c)in the right column. On all graphs horizontal scale represents the tolerance radius from 0% to 100%. Three cases of time-parallel directedness are shown using trajectories (A, B, C, D) from Figure 8: i) Perfect (top graph, trajectories A and B); horizontal *TC* line across the top at 1.0 is characteristic. ii) Complete directional reversion (middle graph, trajectories A and C); much negative deviation from expectation is characteristic. iii) "Chaos" (bottom graph, trajectories A and D); the *TC* graph hogs expectation.

The *TC* values are measures of the strength of topological similarity. The corresponding probability values indicate their significance. A value is considered significant if the chance of any value being at least as large is small under the assumption that the process is ruled by purely random compositional transitions.

f) Since directional scores are used, the order transformation of co-ordinates is implicit. Speaking as if we dealt with process trajectories in real time (trajectory A), the results above point to a perfect relationship of process time and compositional change for trajectory B. Nothing of the sort exists in the trajectory pair A and C. The value 0.5 is random expectation from first principles. The expectation generated in randomisation experiments can differ. The difference of the theoretical expectation and the experimental expectation is bias (see the table above).

g) Up until this point, the tolerance radius was set to 0%. When it is raised to 100%, in 1% increments, graphs are generated for *TC* and related statistics. Figure 9 displays the resulting graphs for the trajectories in Figure 8.

It is to be noted that we may wish to compare an observed trajectory with the "time axis" in order to measure determinism in the observed process (see tabulated text in section on Determinism). In such as case the time axis is visualised as a diagonal line in a reference system whose individual axes are defined by the same vector of time ordinals with elements 1,2,3, …as the co-ordinates. Table 12 includes an example. The computation of *TC* follows the same steps as in any other case.

Appendix C

We are relying on a piece of fact that we took from Mandelbrot (1977). He was interested in the study of irregular and fragmented shapes — like the coastline of Britain, trajectories that we have, Perrin's map of Brownian motion, or a soap flake's perimeter. In order to get a handle on the problem, he had to develop a new analytical geometry almost from the ground up. At the core of what we take from him are the relations

$L(r) = L_0 r^{1-D}$ and $L(r) \sim r^{1-D}$

Symbol L_0 represents the exact length of the perimeter of the measured object, $L(r)$ the length determined at calliper width r, and D a length-related number, called the objects fractal dimension.

When shape is such that no new detail is gained or detail lost under changing calliper width, or in general terms, under changing magnification, theory suggests the value: $D = 1$. It is intuitive that this would be so with our trajectories if the process moved through proportional changes.

At the other extreme stands a trajectory shape generated by pure random transitions. In this case, each step leads to the unexpected, and as theory suggest for this case: $D = 2$. Then D is measuring the shape's increasing complexity (point to point unpredictability when moving with time) from simple 1 to complex 2. How do we determine the value of D in a given case? There is no other way then by some method of approximation. The method we prefer relies on the regression coefficient b which is related to D in the manner of $b=1-D$. The regression line is the best fitting straight line through the point swarm $\{\log L(r), \log r\}$ over a wide range of r values. This means that we measure length and measure it again at different calliper settings, and then we perform a regression analysis.

Given the co-ordinates in Table 13, the corresponding graph is drawn in Figure 10. The exact length of the line

based on the co-ordinates is 99.29 in graph units. But it is not the exact length that interests us most. Our purpose is to derive *D* and for this we need many lengths measured at different calliper settings. The computations follow the Orlóci-Pillar code, available from the authors, emulating the longhand exercise in which a calliper set at a given width is walked over the graph. The results are given in Table 14. Regression line for calliper width increasing to 80 is given in Figure 11.

Our algorithm emulates stepping the calliper through the entire length of the graph until it passes the endpoint The number of complete steps times the calliper width gives the length of the graph $L(r)$ at the given r. The portion at the end of the graph, smaller than the calliper width, is ignored. After the third length measurement, the regression of log $L(r)$ on log r is computed and recomputed at each step as shown in Table 14. Taking calliper widths in steps 1 up to the maximum (90-10) is a decision to have a standard base for comparisons. The question to be put is this: which of the *D* values should be taken as the graph's fractal dimension at the given range of r? The last value of the average in Table 14 is a likely number to be chosen in statistical estimations. We suggest that interested readers subject this suggestion to experimental verification.

The top graph in Figure 12 (based on Table 14) could be used in its entirety. What this graph shows directly is the very complex dependence of *D* on the range of r. The relationship suggests that it should be useful to think of fractal dimension as an entire graph, rather than a single value. But if we wish to choose a single value, we may select the maximum. We find this either as *D* in row 28 or $\overline{D}$ in row 80 of Table 14. The bottom graph in Figure 12 displays the ascending portion of the graphs above it. From this we can see that a linear relationship holds true in the ascending portion. On this basis, we have yet another set of fractal dimensions defined, related to the regression coefficients and the intercepts.

Having clarified the methodology, several observations are in order at this point:

- Before starting the exercise just outlined, the co-ordinate data should really be adjusted to maximum 100 on both axes. Adjustment is necessary to discount the a range related effects.

- As a final example we compare the graph of Figure 10 with the graphs in Figure 13. Two graphs are shown in Figure 13. One is completely regular (top straight line) and the other completely chaotic. Relevant numerical results are given for the straight line and chaotic curve in Table 15. The results suggest

Table 13. Co-ordinate data for illustrations. The total length of corresponding graph is 99.29 in time units.

Time	Entropy
10	10
20	25
30	15
40	17.5
50	20
60	10
70	5
80	5
90	10

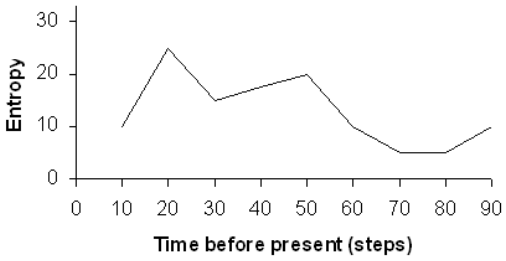


Figure 10. Graph constructed from the co-ordinates in Table 13.

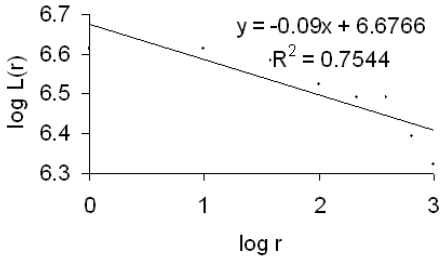


Figure 11. Regression of graph length $L(r)$ on calliper width r. The logarithm is to the base 2. Regression coefficient b= -0.09; fractal dimension D = 1.09. Table 13 presents the data for this graph.

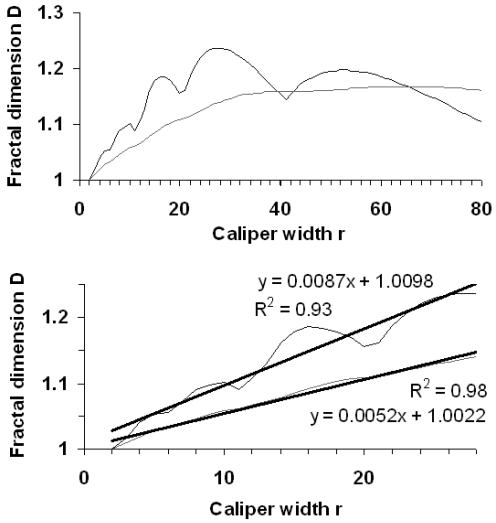


Figure 12. Regression of fractal dimensions on caliper width. See the data in Table 14. Legend: curved line - observed values; smooth line - fitted polynomial.

Table 14. Results of fractal dimension calculations for the graph in Figure 10, based on the data in Table 13. The largest caliper width accommodated is 80 in graph units. True length based on adjusted co-ordinates = 99.288236 in graph units.

Calliper setting r	Length measured $L(r)$	Regression coefficient* b	Fractal dimension D	Cumulative mean fractal dimension $\overline{D}$
1	98			
2	98			
3	96	-0.0167475	1.0167475	1.0083737
4	92	-0.0402675	1.0402675	1.019005
5	90	-0.0529365	1.0529365	1.0274879
6	90	-0.0552176	1.0552176	1.0330338
7	84	-0.0720088	1.0720088	1.0395296
8	80	-0.0899834	1.0899834	1.0467373
...	...	...	...	...
28	56	-0.2371445	1.2371445	1.1398520
...	...	...	...	...
80	80	-0.1039735	1.1039735	1.1617705

* Number of points to which regression line is fitted is the number of lengths measured. This happens to be the same as the row number in which the value of b is entered.

Table 15. Selected results for the graphs in Figures 10 and 13 after adjustments to equal graph size (maximum 100 on all axes).

Regression coefficient b	Fractal dimension D	Cumulative mean fractal dimension $\overline{D}$	Remark
-0.2371445	1.2371445	1.1398520	Graph in Figure 10
-0.07751461	1.0775146	1.0845530	Flat graph, Figure 13
-1.00475150	2.0047515	1.4274527	Graph of random oscillations, Figure 13

two things. First, the fractal dimensions D correspond with theory: 1 for the flat curve and 2 for the random graph. The mean fractal dimension $\overline{D}$ is far below 2. This then suggests that one may do best by using the maximum of D as the fractal dimension in Figure 10, or in general in all cases.

Appendix D

Paleopollen data file posted on the world wide web in 2000 at address http://www.ngdc.noaa.gov/paleo/ftp-pollen.html. Site: Hanging lake, Yukon Territory of Canada. Latitude: 68.23.00. Longitude: 138.23.00. Elevation: 500 m. Contents of data set used: percentages, top 15 pollen types (taxa) + two groups of taxa, 133 paleorelevés. Reliable age bounds: 0 - 33090 yr. Contact person: L.C. Cwyngar. Present vegetation: Taiga.

Paleopollen data file posted on the worldwide web in 2000 at address http://www.ngdc.noaa.gov/paleo/ftp-pollen.html. Site: Lagoa das Patas, Amazon of Brazil. Latitude: 0.16.00N . Longitude: 66.41.00W . Elevation:

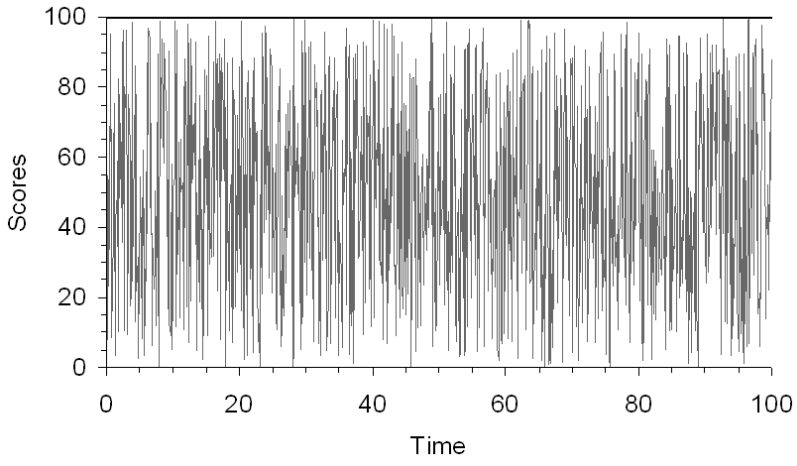


Figure 13. Extreme cases in graph shape. Simple: horizontal line on top. Complex:1000 random oscillations.

300 m. Date of origin: 19 Dec 1997. Contents of data set used: percentages, top 15 pollen types (taxa) + two groups of taxa, 49 paleorelevés. Reliable age bounds: 0 - 42210 yr. Present vegetation: Tropical Rainforest. Contact person: P.E. De Oliveira.