

Journal Article



Multi-scale trajectory analysis: powerful conceptual tool for understanding ecological change



Journal	Frontiers of Biology in China
Publisher	Higher Education Press, co-published with Springer-Verlag GmbH
ISSN	1673-3509 (Print) 1673-3622 (Online)
Issue	Volume 4, Number 2 / June, 2009
Category	Review
DOI	10.1007/s11515-009-0012-y
Pages	158-179
Subject Collection	Biomedical and Life Sciences
SpringerLink Date	Thursday, March 12, 2009

PDF (717.1 KB) Free Preview

László Orlóci^{1, 2}

(1) Ecologia Quantitativa, Universidade Federal do Rio Grande do Sul, Porto Alegre, RS, 91540-000, Brazil

(2) *Present address:* 3-575 McGarrell Pl, London, Canada, N6G 5L3

Received: 5 August 2008 **Accepted:** 5 September 2008 **Published online:** 6 March 2009

Abstract The model at the basis of trajectory analysis is conceptually simple. When applied to time series vegetation data, the projectile becomes a surrogate for vegetation state, the trajectory for the evolving vegetation process, and the properties of the trajectory for the true process characteristics. Notwithstanding its simplicity, the model is well-defined under natural circumstances and easily adapted to serial vegetation data, irrespective of source. As a major advantage, compared to other models that isolate the elementary processes and probe vegetation dynamics for informative regularities on the elementary level, the trajectory model allows us to probe for regularities on the level of the highest process integrity. Theories and a data analytical methodology developed around the trajectory model are outlined, including many numerical examples. A rich list of key references and volumes of supplementary information supplied in the Web Only Appendices rounds out the presentation.

Keywords attractor migration - determinism - fractal dimension - parallelism - periodicity - phase structure

 **László Orlóci**

Email: lorloci@uwo.ca

URL: <http://www.vegetationdynamics.com>

Fulltext Preview (Small, Large)

Multi-scale trajectory analysis: powerful conceptual tool for understanding ecological change

László ORLÓCI^{1,2}

Ecologia Quantitativa, Universidade Federal do Rio Grande do Sul, Porto Alegre, RS, 91540-000, Brazil

© Higher Education Press and Springer-Verlag 2009

Abstract The model at the basis of trajectory analysis is conceptually simple. When applied to time series vegetation data, the projectile becomes a surrogate for vegetation state, the trajectory for the evolving vegetation process, and the properties of the trajectory for the true process characteristics. Notwithstanding its simplicity, the model is well-defined under natural circumstances and easily adapted to serial vegetation data, irrespective of source. As a major advantage, compared to other models that isolate the elementary processes and probe vegetation dynamics for informative regularities on the elementary level, the trajectory model allows us to probe for regularities on the level of the highest process integrity. Theories and a data analytical methodology developed around the trajectory model are outlined, including many numerical examples. A rich list of key references and volumes of supplementary information supplied in the Web Only Appendices rounds out the presentation.

Keywords attractor migration, determinism, fractal dimension, parallelism, periodicity, phase structure

1 Introduction

The field in which this article finds optimal utility is Dynamic Vegetation Ecology. With change in its center of interests, Dynamic Vegetation Ecology is a distant descendent from Kerner von Marilaun's doctrine of vegetation development (Kerner, 1863; Conard, 1951). Since Kerner's time, vegetation dynamics—the phenomenon, causes, and principles of governance—has continued as a central topic in high-level ecological discourse. The theory

and methodology evolved along three lines of reasoning, but in all cases the objectives remained the same—to reveal cause/effect relationships and to identify the principles of process governance. The lines are as follows.

(1) The Kernerian line is that of the thoughtful Naturalist who is reasoning from what the eye can see and letting the mind imagine the rest. This line has led, in Kerner's hands, to the revolutionary finding that vegetation develops and the mechanism involved is facilitation. The thought of vegetation development evolved and became enriched in the succession studies that culminated with Clements (1916) and his contemporaries. McIntosh's (1985) monograph juxtaposes the antagonists and marshals insight from a myriad of references.

(2) Reasoning from surrogate mathematical models about the elementary processes has emerged early in the 20th century as an independent approach. Lotka (1925) and Volterra (1926) are the names credited with the seminal studies. The justification of mathematical models depends on two things: the relevance of the mathematics to the natural process and the dictum that understanding of the mechanisms responsible for the spatio-temporal arrangement of individuals in populations, patches and communities can come from the study of elementary processes such as competition, interaction, and diffusion. The outcome of applications since inception is a mega corpus of highly model-specific conclusions often targeting process stability, occasionally with surprising results (May, 1976, 1981, 1987). Czárán (1998) provides a *tour de force* account of notions and conceptual tools, and Fekete (1985) presents a thematic guide to the mathematically minded student in ecology.

(3) The models above are well-defined as mathematical constructs, such as a system of two quadratic differential equations, a Markov chain of some type, and a Poisson distribution, but remain too simplistic theoretically to serve well the needs of the practicing ecologists. They will do

Received August 5, 2008; accepted September 5, 2008

E-mail: lorlaci@ufrgs.br

1) URL for Web Only Appendix: http://www.springerlink.com/link/Appendices_for_Orlaci
2) Permanent address: 3-575 McGraw-Hill Pl, London, Canada N6G 5L3

References secured to subscribers.

Copyright ©2010, Springer. All Rights Reserved.

MetaPress Privacy Policy

Multi-scale trajectory analysis: powerful conceptual tool for understanding ecological change

László ORLÓCI^{1,2)}

Ecologia Quantitativa, Universidade Federal do Rio Grande do Sul, Porto Alegre, RS, 91540-000, Brazil

© Higher Education Press and Springer-Verlag 2009

Abstract The model at the basis of trajectory analysis is conceptually simple. When applied to time series vegetation data, the projectile becomes a surrogate for vegetation state, the trajectory for the evolving vegetation process, and the properties of the trajectory for the true process characteristics. Notwithstanding its simplicity, the model is well-defined under natural circumstances and easily adapted to serial vegetation data, irrespective of source. As a major advantage, compared to other models that isolate the elementary processes and probe vegetation dynamics for informative regularities on the elementary level, the trajectory model allows us to probe for regularities on the level of the highest process integrity. Theories and a data analytical methodology developed around the trajectory model are outlined, including many numerical examples. A rich list of key references and volumes of supplementary information supplied in the Web Only Appendices rounds out the presentation.

Keywords attractor migration, determinism, fractal dimension, parallelism, periodicity, phase structure

1 Introduction

The field in which this article finds optimal utility is Dynamic Vegetation Ecology. With change in its center of interests, Dynamic Vegetation Ecology is a distant descendent from Kerner von Marilaun's doctrine of vegetation development (Kerner, 1863; Conard, 1951). Since Kerner's time, vegetation dynamics—the phenomenon, causes, and principles of governance—has continued as a central topic in high-level ecological discourse. The theory

and methodology evolved along three lines of reasoning, but in all cases the objectives remained the same—to reveal cause/effect relationships and to identify the principles of process governance. The lines are as follows.

(1) The Kernerian line is that of the thoughtful Naturalist who is reasoning from what the eye can see and letting the mind imagine the rest. This line has led, in Kerner's hands, to the revolutionary finding that vegetation develops and the mechanism involved is facilitation. The thought of vegetation development evolved and became enriched in the succession studies that culminated with Clements (1916) and his contemporaries. McIntosh's (1985) monograph juxtaposes the antagonists and marshals insight from a myriad of references.

(2) Reasoning from surrogate mathematical models about the elementary processes has emerged early in the 20th century as an independent approach. Lotka (1925) and Volterra (1926) are the names credited with the seminal studies. The justification of mathematical models depends on two things: the relevance of the mathematics to the natural process and the dictum that understanding of the mechanisms responsible for the spatio-temporal arrangement of individuals in populations, patches and communities can come from the study of elementary processes such as competition, interaction, and diffusion. The outcome of applications since inception is a mega corpus of highly model-specific conclusions often targeting process stability, occasionally with surprising results (May, 1976, 1981, 1987). Czárán (1998) provides a *tour de force* account of notions and conceptual tools, and Fekete (1985) presents a thematic guide to the mathematically minded student in ecology.

(3) The models above are well-defined as mathematical constructs, such as a system of two quadratic differential equations, a Markov chain of some type, and a Poisson distribution, but remain too simplistic theoretically to serve well the needs of the practicing ecologists. They will do

Received August 5, 2008; accepted September 5, 2008

E-mail: lorloci@uwo.ca

1) URL for Web Only Appendix: <http://www.vegetationdynamics.com> link *Appendices for Ta*

2) Permanent address: 3-575 McGarrell Pl, London, Canada N6G 5L3

much better with a contrasting model type exemplified by Goodall's (1967, 1972) grazing management model that solves a practical problem by employing a carefully put together system of empirical response functions, empirical response rates, and other ecologically sound model elements.

The relevance of any model can be judged by ecologists who understand the mathematics. They usually raise two main objections when it comes to the pure mathematical models. One is the nonlinear nature of biological responses as a rule, and the highly irregular distribution properties of biological variation, also as a rule. These set the ecological conditions as antitheses of the conditions under which a system of linear differential equations and deterministic mathematics like the Poisson and the Markov chain could have much justification (Orlóci, 1993). The other is the impossibility of creating an umbrella theory for natural dynamics based on knowledge of the elementary processes. Gleick (1987) and Çambel (1993) narrate this story at a literate level. Usher (1981, 1992), van Hulst (1992), Anand (1994, 2000), and Anand and Orłóci (1997) make strong points about it in an ecological context. But it is with van Hulst's (2000) paper that the contemplative line in the syndynamic paradigm comes full circle, from reductionist back to holist. The modern arguments bring to mind early readings of the forest ecological classics: Morozov (1912), Cajander (1909), and Sukachev (1913), just to mention a few.

A comprehensive methodology that would serve the specific needs of dynamic vegetation studies on a level comparable in power and completeness to the static multivariate techniques (Greig-Smith, 1952, 1957, 1983; Pielou, 1977; Orłóci, 1978; Legendre and Legendre, 1983; Podani, 1994) is non-existent. The methodology needed should allow the probing of multivariate serial data for deterministic regularities and stochastic effect. There are techniques that take a time series of observations and hand to the user inferences about specific properties of the process that generated the series. Spectral analysis is of this kind for which Reyner (1971) presents the basics in his monograph, Shugart (1977) comments on the method from the point of stand level modelling in forestry, and He and Orłóci (1999) use one of the advanced components, Fourier analysis, to decompose the vegetation time series into additive sine and cosine waves, which can be used to isolate effects specific to factor influences. Other time series techniques focus on testing the efficiency of assumed model processes, such as the Markov chain (Orlóci et al., 1993), as surrogates for the true vegetation process. Still other methods use regression models (Cox and Lewis, 1968) with similar objectives. An interesting turn of mind should be mentioned in which the series is time-static. This is well-presented by Ripley (1981) and Diggle (1981) in their books.

We deal with trajectories, as if each were a continuous uni-dimensional line, turning and twisting in

multidimensional space of the variables that define, as coordinates, the trajectory state at any point in real or model time. This conception of the trajectory implies access to data from the past. Typically, such are the palynological and other types of historic ecological records. The trajectory analysis as we developed is an umbrella methodology under which a broad spectrum of techniques should find specialized niches and feed information concerning different aspects of the vegetation process. Is trajectory analysis data specific? Not really. Once the minimum requirement of a unique temporal or spatial order is met, the data set can be processed. Of course, the specifics of the data will always set limits on what the conclusions can be. What does trajectory analysis do? Simply stated, it captures a likeness of the vegetation process in terms of quantitative variables which are then subjected to scrutiny to reveal information regarding process regularities and governance. Three types of variables are employed.

(1) Variables of the first type fix the elementary compositional transitions and serve as axes for the reference system into which the trajectory is mapped by trajectory analysis. We refer to variables of the first type as primary variables and to the entire reference system based on the primary variables as the phase space of trajectory analysis. The term primary emphasizes the variables' primitive status for being indivisible. This means the primary variables represent the level of detail below which no information is available in the data. Why use the term phase space? Convention calls for it because of the connotation of being simultaneously spatial and temporal.

(2) Variables of the second type are functions of the primary variables in the manner of the compositional transition scalars such as distance, angle, velocity, acceleration, synchronicity scalars (correlation, topological similarity), and disorder and divergence scalars (entropy, information). A detailed account of these is in the succeeding sections where they are used for description of such high-level trajectory properties as phase structure, determinism, periodicity, shape complexity, parallelism, and attractor migration.

(3) Scale variables represent the third type. These fix period length, time step width, order of the scaling functions, units of measurement, and anything else having to do with scales.

What to expect from trajectory analysis? Simply stated, expect a rich stream of information that comes from the data through different channels. The information is always concerned with the parameters of movement and directionality, the nature of forcing factors, and the principles of governance. How does trajectory analysis serve vegetation ecology in the age of global change research? This can be summarized in the following three points.

(1) The service delivered is in the manner of post-dictions and predictions, concerning the properties of vegetation response under dynamic forcing from inside and outside.

(2) Another service has to do with chance. Whatever the actual site or scale, the reign of chance over the vegetation process cannot be complete. In other words, the process will have at least a modicum of determinism, no matter what. Therefore, in post-diction based predictions of the effect of a forcing factor (say global warming) on a vegetation property (say compositional and functional stability), one finds oneself faced with the problem of finding out how much of the total effect is due to chance and how much is due to pure determinism. The isolation of chance effects from determinism is one of the central objectives in trajectory analysis.

(3) The difficulties amplify when a study moves through scales from the global, where determinism is always strong, to the local, where chance effects are most intense. The statistical control of the scale effect is yet another central task performed in trajectory analysis.

This paper has a modular structure within the sections. Each of these takes up a topic and in most cases presents results and a discussion. The first technical section introduces the notion of phase space, the reference frame, and categories of variables fundamental in trajectory analysis. The next section deals with information retrieval by direct probing of the data on a level at which educated impressions are formed about the historic behaviour of individual variables, such as the palynomorphs (taxa) of paleobotany. The following sections go into the analysis and interpretation of complex trajectory properties based on complex trajectory scalars with a multivariate, multi-scale statistical (probabilistic) context. Worked examples, closing remarks, and a rich reference list complete the main text. The text frequently makes use of details published in the Web Only Appendices (web address in the front portion of the paper), full of details and explanations, including many references for new lines to pursue. The source code has been written for specific tasks. An application program package is under preparation.

2 The phase space

This is the space/time reference system of the trajectory model. Considering the perfectly continuous nature of the vegetation process, its perception in the manner of a perfectly continuous trajectory line is appropriate. The only question is whether this line should be one-dimensional (process linear) or multidimensional (process non-linear). The choice is by default the latter since the former is insufficient in all but the most trivial cases treated in Ecology (Orlóci, 1974, 1978, 1993). In fact, the linear model is applicable only and only if the primary variables (taxa, other community elements) respond in a perfectly linear manner to the forcing factors. Fortunately, the linear phase space (not the trajectory line) will cover linear and non-linear cases the same way, but no transformation of the primary variables that are axes is allowed, other than

possibly rigid rotation as happens in Eigenanalysis of the primary variables' covariance matrix as explained in Web Only Appendix A, or conversion to orthogonal functions (Orlóci, 1978) when a one to one correspondence of axes and primary variables is to be retained. In the linear phase space the mapping of a linear trajectory line or any linear segment of the trajectory line will show up as a straight line (which we want) and the mapping of a non-linear trajectory or any non-linear segment of the trajectory as a turning and twisting line (which we also want). Therefore, under no circumstances should the non-linearity be massed with by any of the available techniques (Hill and Gauch, 1980), less intending the distortion of results emitting from trajectory analysis.

The top graph in Fig. 1 portrays an idealized trajectory in three dimensions (X_1 , X_2 , and X_3). In the bottom graph, the phase space is portrayed as a stereogram pair for a real case (Atlantic Heathland, Table 1). Some will require a stereo viewer instrument to see the image rise into three dimensions. In the top graph two lines intertwine, the zigzag line for the natural trajectory, and the dotted line for the best fitting moving Markov chain of the natural trajectory (explained in Web Only Appendix B). The 19 trajectory points (apices in the graphs of Fig. 1) represent the temporal community states (rows in Table 1). The axes defining the bottom graph (not drawn) are Eigenvectors of the species covariance matrix (see Web Only Appendix A). The three stereo dimensions depict the maximum phase space dimensions of the Atlantic Heathland trajectory. The number "3" happens to be the rank of the Atlantic Heathland's covariance matrix. The eight superfluous dimensions are attributed to intensive covariation among the 11 Heathland species (Table 1). When the trajectory is mapped into a reference system of lower than the rank of the full dimensional covariance matrix, loss of information is the result. But the reduction of dimensions is completely under control in two ways: (1) The loss is measurable as the one complement of the fraction of variation (sum for the Eigenvalues, generalized variances) associated with the eliminated dimensions over the total variation (sum of all Eigenvalues or all variances in the original covariance matrix); (2) information loss in variance terms is minimized by eliminating dimensions associated with the Eigenvalues starting with the smallest and going up until a preset threshold is met. Trajectory analysis can, of course, be performed on the trajectory in full dimensions in which case there is no information loss by dimension reduction.

An important component dimension of phase space is time. This is the trajectory mapping itself on which each point is closer to the next point than to the previous one. P is a transition probability matrix (Web Only Appendix B) encapsulating the current vegetation response to intrinsic and extrinsic forcing.

It has already been stated that the purpose of trajectory analysis is to probe the process for regularities. The exercise should bring into play both elementary and

Table 1 De Smidt's Heathland data set after Lippe et al. (1985)

year	BG	EN	CV	ET	MC	CP	JS	RA	OS
1963	57.1	17.9	8.60	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	6.70	55.7	23.3	10.4	0.3	2.0	0.7	0.1	0.7
1970	5.80	61.1	23.7	6.90	0.2	1.2	0.7	0.2	0.3
1971	9.50	57.6	24.7	6.60	0.4	0.6	0.4	0.0	0.3
1972	8.40	62.1	23.7	3.60	0.3	1.2	0.1	0.0	0.6
1973	4.40	67.9	21.3	3.30	0.2	0.6	0.4	0.0	2.0
1974	8.50	58.1	25.8	4.70	0.6	1.3	0.7	0.0	0.4
1975	9.20	62.2	24.3	2.50	0.6	0.9	0.2	0.0	0.1
1976	9.90	58.2	24.9	3.70	0.6	1.1	0.7	0.0	1.0
1977	19.6	48.4	23.5	5.70	0.3	1.2	0.4	0.1	0.9
1978	12.1	58.1	22.7	4.80	0.4	0.4	0.0	0.2	1.3
1979	9.30	65.1	20.3	2.70	0.0	1.5	0.1	0.2	0.9
1980	7.30	68.2	21.5	1.20	0.5	1.0	0.1	0.1	0.2
1981	5.40	65.5	20.8	4.60	1.0	1.6	0.4	0.3	0.6

Site location: 52°N and 6°E. Data elements are point-cover estimates. BG: bare ground; EN: *Empetrum nigrum*; CV: *Calluna vulgaris*; ET: *Erica tetralix*; MC: *Molinia caerulea*; RA: *Rumex acetosella*; JS: *Juncus squarrosus*; CP: *Carex pilulifera*; OS: other species. Significant events: 1964—emergence from period of heavy grazing and fire; 1968: site reaching severe reduction of bare ground; 1976, 1977: draught years.

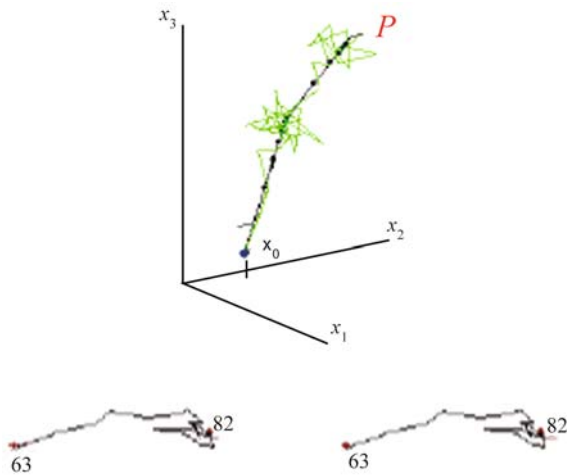


Fig. 1 The phase space diagram after Orlóci et al. (2002a, 2006). Top: idealized trajectory, X_0 to P (the time dimension). Symbols: X_0 : the state, defined by the triplet $(X_{O1}, X_{O2}$ and $X_{O3})$ at which the observer's time begins; P : the state of the momentary target called the attractor. Bottom: stereo mapping of the Atlantic Heathland trajectory (Table 1). See details in the text.

secondary variables. The following text is divided accordingly, starting with the elementary variables. Attention is drawn at this point to the Web Only Appendices C and D for a discussion of important data properties and implications in dealing with the scale effect, and to V. Pillar's MULTIV, O. Wildi's MULVA, and J. Podani's SYN-TAX as relevant application programs to perform

Eigenanalysis and many other data manipulations. Detailed information regarding these programs is published on the internet.

3 Direct probing for regularities in elementary variables

There are time series events that are readily detected on first inspection of the data. New taxa entering the records or others dropping out or change dominance are of this kind. There are other events whose detection requires specialized analytical techniques. One of the techniques uses deviation graphs. The deviations are taken from random expectation according to $\Delta_{ij} = f_{ij} - f_{ij}^o$ and $f_{ij}^o = \frac{f_{i..} f_{.j.}}{f_{..}}$. In the expressions, i identifies a taxon and j a time step. Symbol f_{ij} stands for the quantity of taxon i at time step j . A dot replacing a subscript indicates summation over the subscript replaced. Δ_{ij} may be positive or negative if not zero.

The Hanging Lake spectrum (Table 2, Fig. 2) is considered in further detail. Figure 3 has deviation graphs for selected taxa. The first of the 8 plates has two graphs, T for Vostok temperature differences on the top and sums of squared deviations (SSD) (lower graph) for the sum of squared deviations. The SSD graph is a simple sum of the squared deviations of the individual taxa. Note that the extreme SSD peaks occur when directedness is maximal.

Table 2 Data sources by location and contact person after Orlóci et al. (2006)

location and contact person	latitude longitude altitude /m	number of taxa number of time steps	period covered yr BP	regional vegetation formation annual precipitation/cm potential evapotranspiration/cm
(1) Lagoa das Patas, Amazonas— P. E. Oliveira (Colinvaux et al., 1996)	00.16.00 N 66.41.00 W 300	179 49	0–44 569	TR > 200 120–160
(2) Joe Lake, Alaska— P. M. Anderson (1988)	66.46.00 N 157.13.00 W 183	90 87	0–43 804	T,TA 25–50 < 40
(3) Camel Lake, Florida— E.C. Grimm (Watts et al., 1992)	30.16.00 N 85.01.00 W 20	147 116	0–36 658	LSC 100–150 80–120
(4) Hanging Lake, Yukon Territory— L. C. Cwynar (1982)	68.28 N 138.23 W 500m	89 133	0–41 134	T 25–50 < 40
(5) Jack London Lake, Magadan Oblast, Russia— P. M. Anderson (Lozhkin et al., 1993)	62.10.00 N 149.30.00 E 820	72 60	221–29 876	AT,TA 25–50 < 40
(6) Jackson Pond, Kentucky— G. R. Wilkins (Wilkins et al., 1991)	37.27.00 N 85.43.00 W 212	71 58	0–20 477	TDF 100–150 80–120
(7) Cambará, Rio Grande do Sul, Brazil— H. Behling (Behling et al., 2004)	29.03.09 S 50.06.04 W 1046	164 190	0–42 784	ARA,G 150–200 80–120
(8) Lake Patzcuaro, Michoacán de Ocampo, Mexico— W. A. Watts (Watts and Bradbury, 1982)	19.35.00 N 101.35.00 W 2044	53 64	20–44 100	XF 50–100 120–160
(9) Rusaka Swamp, Burundi— R. Bonnefille (Bonnefille et al., 1995)	03.26.00 S 29.37.00 E 2070	179 141	796–11 910 (46 666)	TG,TH 50–150 120–160
(10) Lynch's Crater, Queensland, Australia— A. P. Kershaw (1994)	17.22.00 S 145.42.00 E 760	22 44 (234)	868–40 000 (–192 649)	EAS 50–150 120–160
(11) Harberton, Tierra del Fuego— V. Markgraph (no reference given).	54.53.00 S 67.10.00 W 20	33 81	0–13360	DS 25–50 80–120
(12) Lake George, NSW— G. S. Hope (Singh and Geissler, 1985)	35.05.00 S 149.25.00 E 673	93 30 (68)	1–40 000 (116 711)	EASC 25–100 80–120
(13) Potato Lake, Arizona— R. S. Anderson (1993)	34.27.43 N 111.20.43 W 2205	77 61	1389–35 271	MDSC 25–50 80–160
(14) Hay Lake, Arizona— B. F. Jacobs (1985)	34.00.00 N 109.25.30 W 2780	44 46	106–44 692	MDSC 25–50 80–160

The number of taxa shown is the number listed in the original records. This is not necessarily the same number as actually used in the analysis, after exclusion of ambiguous entries. Palynological data were downloaded from the 2007 address*. Vegetation classification and precipitation data follow Kühler (1990) and Trewartha (1990). Potential evapotranspiration values are after Trewartha (2001). Abbreviations: T: tundra; TR: tropical rainforest; LSC: lowland shrub conifer; TA: Taiga; AT: Alpine tundra; TDF: temperate deciduous forest; ARA: Araucaria forest; TG: tropical grassland; G: other grassland; XF: xerophytic forest; TH: thorn shrub; EAS: eucalyptus, Acacia shrub; EASC: EAS plus conifer; MDSC: Montane desert shrub conifer; DS: desert shrub.

*<http://www.ncdc.noaa.gov/paleo/pollen.html>

Note also that the flat and low SSD graph in one portion of the time scale is still before 19 kyr BP. A flat graph testifies to low intensity transitions close to random transitions. Sharp oscillations begin after 19 kyr BP. This is heralding the advent of intense dynamics and increased vegetation

instability coinciding with the period of sustained climate warming that culminates around 11 kyr BP.

Considering the individual graphs in the deviation (Δ) graphs in the second and subsequent plates of Fig. 3, we note the exceptionally sharp deviation peaks for, for

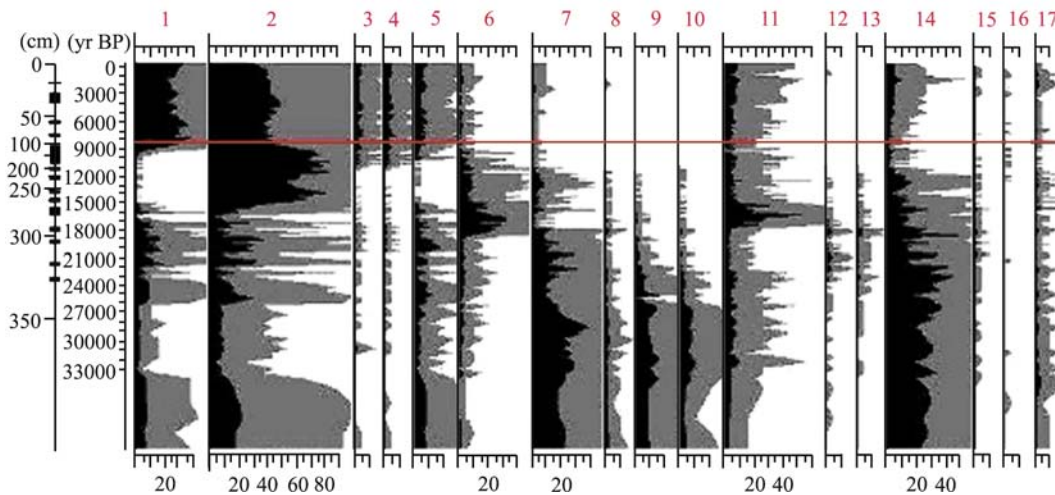


Fig. 2 Cwynar's (1982) palynological spectrum from Hanging Lake, Yukon Territory. Sample and location information are given in Table 2. Graph and accompanied data were downloaded from the Global Pollen Database, 2007 URL address: <http://www.ncdc.noaa.gov/paleo/pollen.html> and search. Short listed taxa (each taxon represents a different palynomorph named according to their lowest identifiable systematic status) after L. C. Cwynar: 1, *Alnus*; 2, *Betula*; 3, ericaceae; 4, Ericales; 5, *Picea*; 6, *Salix*; 7, *Artemisia*; 8, Asteraceae-Asteroidae; 9, Brassicaceae; 10, Chenopodiaceae-Amaranthaceae; 11, Cyperaceae; 12, Fabaceae; 13, *Plantago canescens*; 14, Poaceae; 15, Rosaceae; 16, Other trees and shrubs; 17, other herbs; bottom scale, pollen counts (%); dark shading, original scale; light shading, 5×exaggerated scale. Black markings on depth scale: dated horizons. Red line demarks a record set called *paleorelevé*. This is the virtual equivalent of the regional vegetation community 8200 years BP. Interesting to observe the great age of the sediments. From this we can infer that the Hanging Lake site remained essentially undisturbed by glacial ice for at least forty millennia.

example, the gonidium (Pe) and asco (Fu) spore types in the middle period. These indicate in all likelihood the lichen element reaching maximum representation in the region. The Fu and Pe graphs are in stark contrast with the cold steppe taxa, such as the Poaceae (Po), *Artemisia* (Ar), *Chenopodium* (Ch), and Brassicaceae (Br). Each of these outperforms random expectation in the early, cold period. Interpretation of other graphs follows along similar lines. Low graph segments close to random expectation indicate high stability, and the intensively oscillating peaks up or down from random expectation signal high instability. Some interesting facts are found in Table 3. Periods, characteristic taxa, and vegetation formation types are listed. Clearly seen from this, the effect of Holocene climate warming and concomitant humidity increase species transitions. A most remarkable indicator of this is the peaking of *Sphagnum* (Sp) around 5000 BP. Peat bog formation must have reached maximum extent around that time.

4 Complex trajectory properties—a preview

These properties are linked to the behaviour of the secondary variables. Six complex properties are presented and later discussed in detail.

(1) Determinism. This is the tendency of the vegetation process not to represent a pure case in random walk through time. It has to be emphasized even at the risk of being repetitious that in the real world neither pure

determinism nor pure randomness is to be expected, but an intricate convolution of the two. This prevents the trajectory from being aimed at a stationary attractor P or being completely aimless. The convolution of determinism and randomness creates an oscillating attractor $P = \pi + \delta$, which, as an uncertain target with uncertainty proportional to δ , forces the trajectory to re-aim itself continuously.

(2) Phase structure. Given sufficient time, the vegetation process will go through phases. The stereogram in Fig. 1 illustrates a rather clearly defined phase structure at the scale of annual recording: a fast moving linear phase followed by a rather convoluted non-linear phase. The first phase is manifesting non-equilibrium dynamics (Bartha et al., 1998), almost as strikingly linear as in the case of the stationary Markov process. The first phase lasts until about 1970, and then the second phase takes it over. This is a turbulent process phase with the imprints of high stability at the scale of the observations. Inspection of Table 1 reveals the cause that triggered phase change: the community is reaching the site's carrying capacity.

(3) Periodicity. This property is illustrated in Fig. 4. Compositional transition velocity V (the amount of change in unit time), acceleration A , compositional distance d and the acute angle AN enclosed by successive trajectory segments are the scaling functions. The next section gives the functional forms. The graphs prove what is expected of Natural periodicity: high irregularity for both amplitude and wavelength.

(4) Shape complexity. This term is used in Mandelbrot's (1977) fractal geometric sense (Web Only Appendix E).

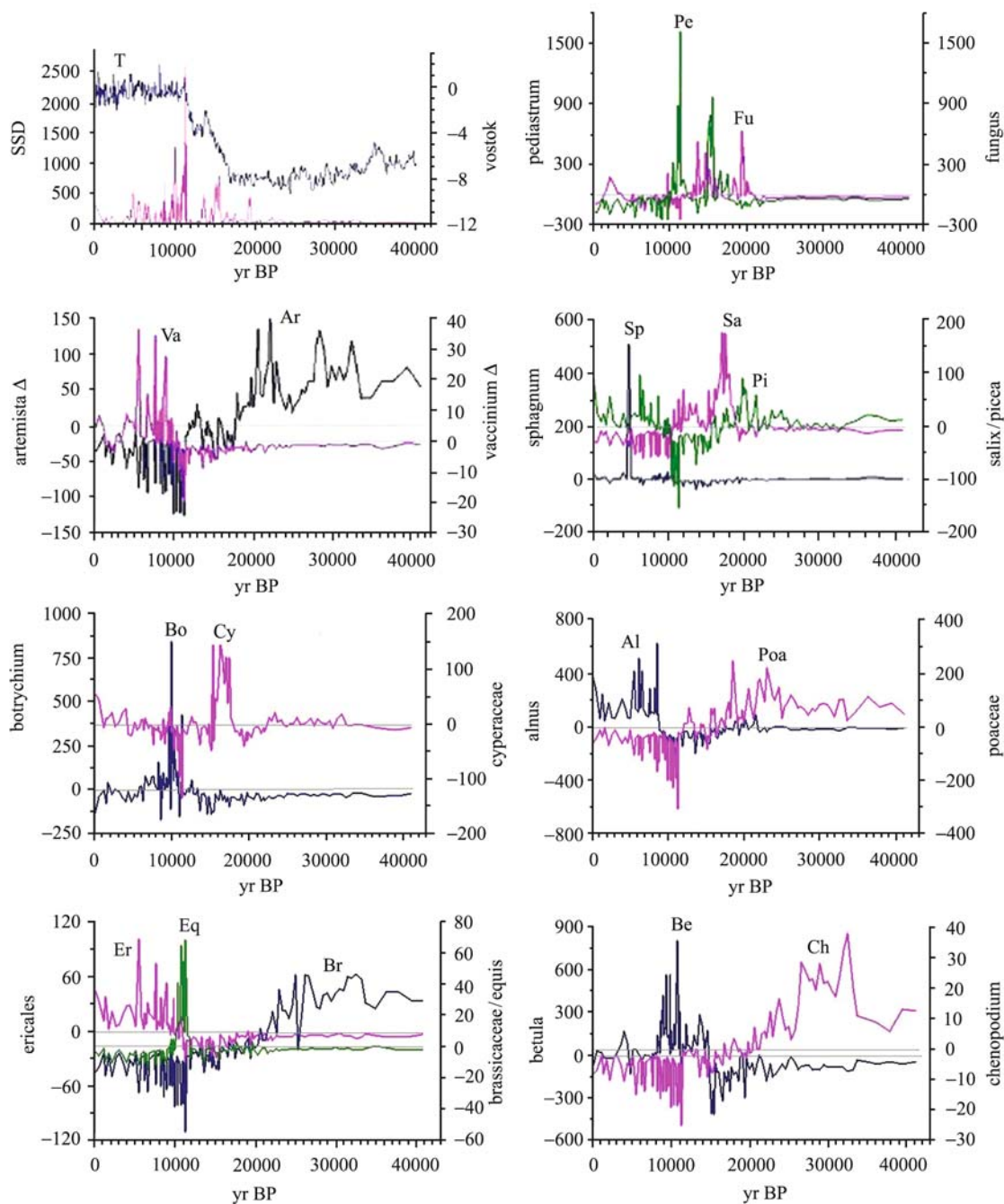


Fig. 3 Long-term performance of taxa in deviation terms after Orlóci et al. (2006). Taxa are taken from Cwynar's (1982) Hanging Lake spectrum (Fig. 2, Table 2). Horizontal axis: years before the present; vertical axis: Vostok temperature differences (T) following Petit et al. (2001); SSD: sums of squared deviations; Δ : deviations from random expectation. Random expectation is the imaginary state that we would have if the compositional transitions were ruled completely by chance. Deviations above or below random expectation (the zero line) indicate over or under-performance of the taxon. See Feoli and Orlóci (1985) and Orlóci and Orlóci (1988) for use of Δ and SSD graphs in a study of environmental effects in transect based edge detection.

According to this, shape complexity of the process trajectory is minimal in the linear phase and maximal in the non-linear phase.

(5) Parallelism. This is a surrogate term for process

co-ordination in different sites (Web Only Appendix F). The measurement of parallelism has considerable importance for interpretation of the long-term process as a landscape or regional pattern. The idea of process

Table 3 Late Quaternary vegetation history of Hanging Lake reconstructed from the sums of squared deviations (SSD) graph (Fig. 3)

time period kyr BP	indicator taxa	formation type
42—19	Poaceae, <i>Artemisia</i> , Brassicaceae, <i>Chenopodiaceae/Amaranthaceae</i>	dwarf shrub steppe of windswept dry uplands
19—11	<i>Pediastrum</i> (alga), Fungi, <i>Salix</i> , Cyperaceae, <i>Picea</i>	lichen dominated tundra, wetlands expanding
11—0	<i>Betula</i> , <i>Alnus incana</i> , <i>Bothrichium</i> , <i>Sphagnum</i> , <i>Picea</i> , Ericales, <i>Equisetum</i> , <i>Vaccinium</i>	taiga uplands and wetland mosaic

See site details in Table 2.

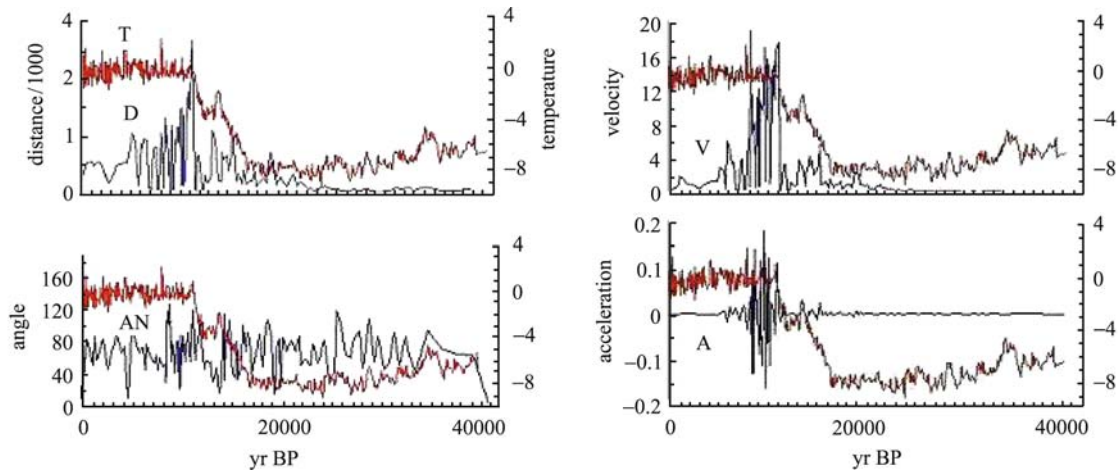


Fig. 4 Long-term oscillations of distance (D), angle (AN), velocity (V) and acceleration (A) at the Hanging Lake site (Table 2). The Vostok temperature graph (T) is superimposed. The functions are discussed in the text. Data set used covers 89 taxa. See Petit et al. (2001) for temperature details. Diagram adapted from Orłóci et al. (2006).

parallelism already occurred to von Post more than a half century ago (von Post, 1946) and led him to a simple graphical solution of its measurement based on palynological spectra. The technique to be reviewed is more advanced, probabilistic, multi-scale and suitable for computer processing

(6) Attractor migration. The ever oscillating target P has far-reaching consequences in requiring a continuous re-aiming of the process. The re-aiming for whatever reason is manifested by the trajectory points not falling on a straight line in the same direction. In general, the lesser the inside angle at an apex is, the more pronounced the change in P .

Emphasis on the mentioned trajectory characteristics is not exclusive. Others will be introduced later. But before continuing with the topic, the scaling functions have to be defined.

5 Functions for complex trajectory scaling

The trajectory properties to be scaled are compositional and functional structures, the products of complex vegetation responses to internal and external forcing. Transition, synchronicity, disorder, and divergence

functions are considered in detail. Possible others are left for the readers to discover.

5.1 Compositional transition functions

Euclidean distance of two paleorelevés Δt time apart:

$$d_t = \sqrt{\sum_{i=1}^p (f_{it+\Delta t} - f_{it})^2}.$$

The relevé that describes a past state is called *paleorelevé*. Regarding the time points, t and $t+\Delta t$ represent the natural order of progression from the past to the present or from horizons lower to horizons higher in the sediment core. Symbol i is a label for taxa and p for the number of taxa.

Acute angle α_B enclosed by trajectory segments A to B and B to C:

$$d_{AC}^2 = d_{AB}^2 + d_{BC}^2 - 2d_{AB}d_{BC}\cos\alpha_B.$$

The values of α_B range from 0° to 180° . A zero value indicates a complete about-face of the process at point B . At $\alpha_B=180^\circ$ the trajectory continues in the same track. In general, the sharper the angle, the more extreme is the

directional change, and in turn, the more unstable is the attractor. The segment length from *A* to *B* or from *B* to *C* is proportional to the intensity of change.

Transition velocity: $V = \frac{d}{\Delta t}$. This is the average compositional transition velocity within the Δt time interval. Since d_t is compositional distance, V is always positive. It can be interpreted as the level of compositional instability within the Δt time interval.

Acceleration, deceleration: $A_t = \frac{V_{t+\Delta t} - V_t}{\Delta t}$. A_t is directional, positive (acceleration) or negative (deceleration). Extreme peaks in the A_t graph identify time points of dramatic compositional transitions, the so called hotspots of change.

Sample graphs are shown for the Hanging Lake spectrum in Fig. 4. These and the graphs in Fig. 3 reinforce the site history as narrated in Table 3. These underline also the specific message of change: flat graph segments indicate periods of stability. Stability and instability are relative and their perception is dependent on the time scale. Furthermore, there cannot be stagnation; that is, the ratio of stability/instability is ever changing. The sharp high amplitude D , V and A peaks are to be focused in the interpretation of the graphs, for these peaks are indicators of extreme vegetation dynamics and levels of compositional stability/instability at the hot spots of change.

5.2 Synchronicity functions

Correlation:

$$r(V,T) = \frac{\sum_{i=1}^n (V_i - \bar{V})(T_i - \bar{T})}{\sqrt{\sum_{i=1}^n (V_i - \bar{V})^2 \sum_{i=1}^n (T_i - \bar{T})^2}}$$

As written, the symbols indicate that correlation of oscillations in transition velocity V and atmospheric Vostok temperature T is measured. Symbols \bar{V} and \bar{T} are

average values over n time steps. The $r(V,T)$ scalar is best not to be computed directly from the V and T values (as taken from the graphs such as in Fig. 4). They should rather be based on residual values after error dampening. The method of error dampening recommended is based on block averages as shown by examples in Table 4 and regression estimation as in Fig. 5. An essential aspect of the technique is scaling by increasing the time step width $BS = 1, 2, \dots, M$ and calculating a new $r(V,T)$ at each BS . The question arises on which of the M values of $r(V, T)$ should serve as the estimate $\rho(V,T)$. We opt for the regression estimate corresponding to the original step size $BS = 1$ in the manner of Fig. 5 and Table 5.

Regression analysis gives us estimates for ϕ^+ and ϕ^- based on the F^+ and F^- graphs in Fig. 5 that also includes the $r(V,T)$ graph corresponding to increasing step size $BS = 1, 2, \dots, M$. Regression analysis is performed on the $r(V,T)$ graph. The regression estimate at $BS = 1$, symbolically $\rho(V,T)$, is the final estimate of synchronicity. How do we get the F^+ and F^- values? A large number of randomly sited substrings of random length are selected from the V and T series at each BS (no substring shorter than 5 consecutive time steps) and for each substring a new $r(V, T)$ is calculated. The results are a large number of BS -specific $r(V, T)$ values in which the positive and negative cases are counted as F^+ and F^- . On the F^+ and F^- graphs we perform a regression analysis to obtain the frequency estimates ϕ^+ and ϕ^- at $BS = 1$.

How do we interpret Table 5? Clearly, what matters most is the magnitudes of $\rho(V,T)$, ϕ^+ and ϕ^- by which the synchronicity of the V and T strings can be estimated. Strong positive synchronicity is indicated when a positive $\rho(V,T)$ is numerically large. Formation specificity of the synchronicity is indicated by the dominant ϕ^+ or ϕ^- . On this basis, Table 5 supports the conjecture put forward earlier (Orlóci et al., 2002a, b, 2006) that the dominance of ϕ^+ or ϕ^- is correlated with regional aridity. In humid regions, ϕ^+ is dominant. In the arid regions, ϕ^- is the dominant type. What do we conclude? Process acceleration is expected under global climate warming.

Table 4 Error dampening by moving averages

block size (BS)	variable	values of V and T at time points before the present										$r(V,T)$
		1	2	3	4	5	6	7	8	9	10	
1	V	3	3	5	6	6	1	4	8	3	9	0.963
1	T	-1	-2	3	4	4	-3	1	5	0	6	
2	V	3	4	5.5	6	3.5	2.5	6	5.5	6		0.963
2	T	-1.5	0.5	3.5	4	0.5	-1	3	2.5	3		
3	V	3.7	4.7	5.7	4.3	3.7	4.3	5	6.7			0.996
3	T	0	1.7	3.7	1.7	0.7	1	2	3.7			
4	V	4.3	5	4.5	4.3	4.8	4	6				0.934
4	T	1	2.2	2	1.5	1.8	0.8	3				

Given the trajectory's length in n time steps, the number of moving averages at block size BS will be $n - BS + 1$. Block size is measured in initial time step units 1, 2, 3, ..., $M < n$. In the example below, $n = 10$ and $M = 4$. M is the chosen upper limit for BS . The synchronicity measure in the last column is a product moment correlation coefficient.

Table 5 Regression estimates $\rho(V,T)$, ϕ^+ and ϕ^- in 14 sites

locality	$\rho(V,T)$	LL	UL	R^2	ϕ^+	ϕ^-	CVF	Thi
Lagoa das Patas 00.16°N 6.41°W	0.12	-0.01	0.28	0.73	80.96	16.30	TR	> 1.43
Joe Lake 66.46°N 157.13°W	0.35	0.30	0.40	0.98	80.90	16.76	T	1.88
Camel Lake 30.16°N 85.01°W	0.31	0.27	0.35	0.98	75.84	20.68	LSC	1.25
Hanging Lake 68.28°N 138.23°W	0.63	0.56	0.71	0.91	68.68	30.22	T TA	1.88
Jack London L. 62.10°N 149.30°E	0.11	0.02	0.20	0.98	65.08	29.16	AT TA	1.88
Jackson Pond 37.27°N 85.43°W	0.30	0.28	0.32	0.98	57.31	38.39	TDF	1.25
Cambará 29.03°S 50.06°W	0.53	0.50	0.56	0.95	56.63	32.77	ARA G	1.75
Lake Patzcuaro 19.35°N 101.35°W	0.60	0.50	0.70	0.78	52.79	41.25	XF G	0.54
Rusaka Swamp 3.25°S 29.37°E	0.23	0.17	0.24	0.86	46.84	39.99	TG TH	0.71
Lynch's Crater* 17.22°S 145.42°E	-0.66	-0.70	-0.62	0.74	42.62	55.08	EAS	0.71
Tierra del Fuego 54.53°S 67.10°W	-0.25	-0.37	-0.15	0.55	33.64	63.71	DS	0.36
Lake George* 35.05°S 149.25°E	-0.48	-0.99	-0.28	0.57	15.12	84.32	EASC	0.63
Potato Lake 37.27°N 111.20°W	-0.31	-0.42	-0.19	0.86	11.15	84.91	MDSC	0.31
Hay Lake 34.00°N 109.25°W	-0.23	-0.34	-0.12	0.85	4.74	93.91	MDSC	0.31
mean (positive)	0.35	0.29	0.42	0.91	65.00	29.50		1.39
mean (negative)	-0.39	-0.56	-0.27	0.71	21.45	76.39		0.46
grand mean	0.09	-0.02	0.17	0.84	49.45	46.25		1.04

Estimation technique (Fig. 5) and error dampening are multi-scale (Table 4), after Orlóci et al. (2002a, 2006). Abbreviations: V , compositional transition velocity; T , Vostok temperature differences; $\rho(V,T)$, regression estimate of $r(V,T)$ at $BS=1$; BS , block size; R^2 , coefficient of determination; ϕ^+ , regression estimate of the frequency F^+ of positive correlation values at $BS=1$; ϕ^- , regression estimate of frequencies F^- of negative correlation values at $BS=1$; LL and UL: lower and upper limits of the 95% confidence interval about the regression line (normally distributed statistical errors assumed); Thi, Thornthwaite index (precipitation per potential evapotranspiration); CVF, current regional vegetation formation; T, Tundra; TR, Tropical rainforest; LSC: Lowland shrub conifer; TA, Taiga; AT, Alpine tundra; TDF: Temperate deciduous forest; ARA: Araucaria forest; G: temperate grassland; TG: Tropical grassland; XF: Xerophytic forest; TH: Thorn shrub; EAS: Eucalyptus, Acacia shrub type; EASC, EAS plus conifer; MDSC, Montane desert shrub conifer; DS, Desert shrub.

Some correlation values: $r(\rho(V, T) \times Thi) = 0.621$, $r(\phi^+ \times Thi) = 0.821$, $r(\phi^- \times Thi) = -0.804$.

*Current 40 kyr period analyzed.

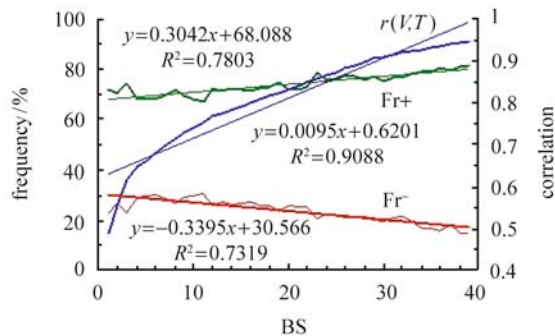


Fig. 5 Regression method of estimation using block size (BS) as the scale variable. The site is Hanging Lake (Table 2, Fig. 4). See numerical results in Table 5. Abbreviations: F^+ or F^- , percent frequency of positive or negative correlation values determined in randomized sitting of windows with randomly chosen BS values; y : regression estimate; R^2 , coefficient of determination; BS : block size in time step units; $r(V,T)$, product moment correlation coefficient. The 'best estimate' used is the point on the regression line corresponding to $BS=1$. Graph adapted from Orlóci et al. (2002a, 2006).

Until this point, a time equivalent pairing of the V and T values has been retained. But what will happen with $\rho(V, T)$, ϕ^+ and ϕ^- if the temperature series is shifted backwards in time (right in the V, T graphs) so that the time lag of the V and T values paired widens from 0 to 1

then to 2, 3 and so forth in, say, millennial steps. At each time lag the $\rho(V, T)$, ϕ^+ and ϕ^- estimates are recomputed for $BS=1$ in the same manner as already described. Orlóci et al. (2006) discuss this case at length and present examples.

5.3 Topological similarity coefficient

The only way to generalize the trajectory as a geographic pattern is by a collection of time series from different sites. A structure can be defined on the collection by some similarity scalar matrix whose elements measure time series parallelism in pairs. Since neither the time points nor the list of taxa may be comparable as a general case, the techniques so far discussed may not be satisfactory. The synchronicity scaling function to be discussed is a topological similarity (TC in Web Only Appendix F) which generates scale-dependent similarity values. For an original description and overview, readers are referred to Orlóci (in Fekete et al., 1998; Orlóci et al., 2002a, 2006).

TC makes provision for the effect of excessive random variation by defining an expandable tolerance sphere about the trajectory points. Any trajectory point found within a tolerance sphere is considered identical to the point in its center. The radius of the tolerance sphere is increased in steps to determine the amount of variation to be removed

Table 6 Characteristic values of the topological similarity coefficient and related statistics

trajectory pair (columns) and statistics (rows)	Hanging Lake × Cambará	Hanging Lake × Lagoa das Patas	Cambará × Lagoa das Patas	Lagoa das Patas × RND	identicals	RND × RND
observed topological similarity at maximum deviation from random expected	0.8991	0.8260	0.8701	0.7019	1.0000	0.3438
random expectation (bias) at maximum deviation	0.3861	0.3861	0.4196	0.5449	0.3407	0.3355
deviation from expectation (unbiased similarity)	0.513	0.4398	0.4504	0.1569	0.6593	0.0083
tolerance radius % at maximum deviation	60	50	56	71	35	36
lower 95% confidence limit	0.2813	0.2813	0.3167	0.5111	0.2105	0.3123
upper 95% confidence limit	0.4909	0.4909	0.5225	0.5788	0.4709	0.3586
significance	high	high	high	high	very high	none

See graphs in Fig. 6, localities in Table 2, and method description in Web Based Appendix F. RND indicates a trajectory, emulating chance driven compositional transitions. The expected values, positive or zero, can be considered as measures of bias in the similarity coefficient at the given tolerance radius. Therefore, the observed similarity's deviation from their random expectation expresses the trajectories' level of true parallelism. Expected values and confidence limits are determined in Monte Carlo (randomization) experiments (Metropolis and Ulan, 1949; Hammersley and Handscom, 1967; Metropolis, 1987; Edgington, 1987; Pillar and Orlóci, 1996).

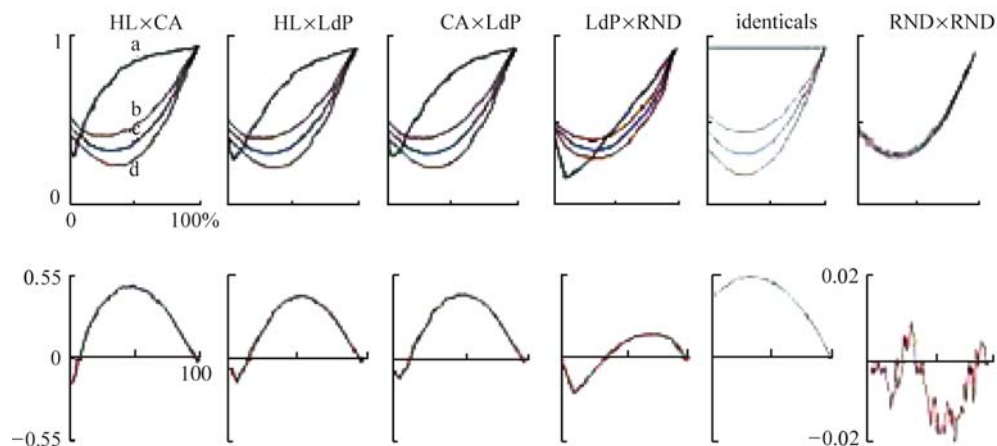


Fig. 6 Top row: graphs of the topological similarity index and the 95% statistical confidence limits; bottom row: the similarity coefficient's deviation from random expectation determined in Monte Carlo experiments; horizontal scale: the radius of the tolerance sphere (scale variable) extending from 0 to 100% in 1% steps. See method description and references in Web Based Appendix F and numerical results in Table 6. Data transformation to equal time steps is applied (see Web Based Appendix H). Abbreviations: a, topological similarity; b, upper limit of 95% (statistical) confidence interval about random expectation; c, expectation of *a* under assumption of random compositional transitions; d, lower limit of 95% confidence interval. Localities are identified in Table 2: HL, Hanging Lake, Yukon Territory; CA, Cambará, Rio Grande do Sul; RND, series of random numbers emulating the case of chance driven compositional transitions; identicals, a trajectory and its exact copy. We take as an optimal estimate of parallelism the value of the similarity coefficient at maximum deviation from random expectation. To attain this value about 50% of random variations have to be bridged by the tolerance radius. Any similarity value (points on curve 'a') outside the confidence limits is deemed significant at better than 2.5%. Graph adapted from Orlóci et al. (2006).

before parallelism could be seen at its best. When no trace of parallelism is present other than that expected by chance, the similarity graph takes the shape of cases LdP×RND and RND×RND in Fig. 6 and Table 6. The

similarity graph becomes a horizontal line at $TC = 1$ when the trajectories are identical.

When the co-ordinates for both trajectories are long chains of random numbers, the topographic coefficient's

expectation comes close to 0.5 at 0% tolerance radius. The expected value tends to decline at first and then increases again with the broadening tolerance radius (graph ‘c’ in Fig. 6). The expected value at any tolerance radius is a threshold for comparison. Any point on the ‘a’ graph (Fig. 6) higher than the expected value “c” indicates co-ordination higher than that expected by chance. Where the point on ‘a’ is lower than on “c”, the trajectories are deemed less co-ordinated than would be expected by chance. The shape of graph ‘a’ is indicative of the intensity of parallelism over the range of the tolerance radius. The more convex the shape is, the greater the maximum co-ordination. By this criterion, the LdP and RND trajectories of Fig. 6 have very little in common, but HL and CA have very high similarity. Portions of the ‘a’ curve outside the 95% confidence limits indicate statistical significance at better than 0.025 probability.

Once the two trajectories’ co-ordination is declared significant, the trajectories are deemed parallel, and attention turns to the deviation graphs (2nd row in Fig. 6) to find the tolerance radius at which co-ordination comes to its most intense. It should be noted that the further one moves from the zero point on the tolerance radius axis, the more variation is removed. First, the appearance of co-ordination increases, and then it declines. As a general tendency, the step-by-step broadening of the tolerance radius encourages first a steady rise in the deviation graphs, then a steady decline. The inflexion point appears to be drifting to higher and higher values as the similarity curve lowers itself and becomes more and more concave. At the extreme where the reign of chance is complete, the deviations graph shows multi-modal oscillations across the zero line within the very narrow confidence limits (Fig. 6).

5.4 Disorder and divergence functions

Rényi’s (1961) disorder related entropy of order α : $H_\alpha = \frac{1}{1-\alpha} \ln \sum_{i=1}^s p_i^\alpha$. This is defined for any s -valued distribution $\mathbf{P}=(p_1 p_2 \dots p_s)$. In the cases to be discussed, i selects one of the s taxa, $p_i = \frac{f_i}{T}$, f_i is a palynological particle count, and $T = \sum_{i=1}^s f_i$. Symbol α is a scale variable

whose values may range from 0 up to any positive number, but never exactly 1. Depending on α , different types of ecological diversity indices are defined. For example, for $\alpha = 0, 1, 2$, H_α defines richness, Shannon’s entropy, or a form of the Simpson index. We refer for more details to Orlóci (1991a,b,c) and Web Only Appendix G.

Rényi’s (1961) information divergence of order α : $I_\alpha = \frac{1}{\alpha-1} \ln \sum_{i=1}^s \frac{p_i^\alpha}{q_i^{\alpha-1}}$. The divergence of the s -valued observed distribution \mathbf{P} is expressed in information terms

from an identical valued standard distribution \mathbf{Q} that is given a priori. For example, \mathbf{Q} may be a maximally even distribution in which all the elements are equal to $1/s$. \mathbf{P} and \mathbf{Q} are such that $\sum_{i=1}^s p_i = \sum_{i=1}^s q_i$. H_α and I_α generate curves as functions of α (the scale variable). To find the points H_1 and I_1 , set α equal to a value close to 1, say 0.99999, or use the alternative forms $H_1 = -\sum_{i=1}^s p_i \ln p_i$ and $I_1 = \sum_{i=1}^s p_i \ln \frac{p_i}{q_i}$.

Oscillograms of the long-term evolution of H_1 and I_1 are shown for the Hanging Lake site in Fig. 7. Table 7 includes regression estimates of the correlation of H_1 and I_1 with the Vostok temperature series T for $BS=1$ according to Fig. 8. The technique of estimation is the same as the one discussed in connection with Fig. 5.

Table 7 Regression estimates of entropy and information in the Hanging Lake site (Table 2) for $BS=1$ corresponding to Figs. 7 and 8

	r	F^+	F^-
entropy	-0.5236	25.4281	70.7639
information	0.7141	69.7587	28.2512

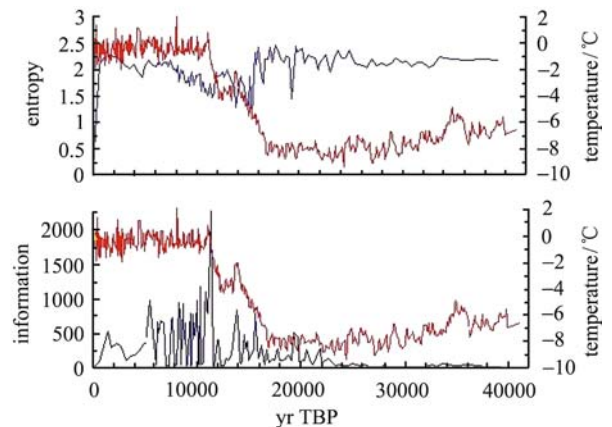


Fig. 7 Long-term evolution of taxon diversity (H) and information divergence (I) at the Hanging Lake site. Refer to Rényi (1961) and Web Based Appendix G for technique. Vostok temperature graph follows Petit at al. (2001). Site detail is in Table 2.

6 Further on complex trajectory properties

6.1 Phase structure

Short and long-term, directed and random attractor migration generates the trajectory’s phase structure. The trajectory’s mapping indicates this in phase space by the changing intensity and frequency of the zigzagging movement. For example, a single major attractor event in

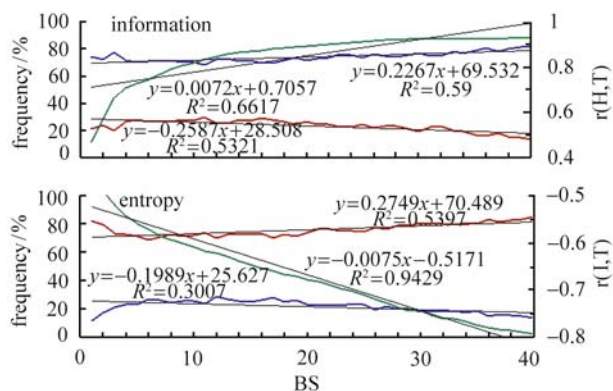


Fig. 8 Regression estimation of the correlation of entropy (H) and information divergence (I) with Vostok temperature (T), and the frequency distribution of the correlation as a function of block size BS at the Hanging Lake site (Table 2). See oscillograms in Fig. 7. F^+ or F^- : percent frequency of positive or negative correlation values; y : regression estimate; R^2 : coefficient of determination; BS: block size in time step units; $r(H,T)$ or $r(I,T)$: product moment correlation coefficient. Regression estimates at BS = 1 are given in Table 7. Note the negative correlation coefficient for entropy and the positive correlation for information divergence. These support the hypothesis that climate warming destabilizes the vegetation.

Table 8 Phase structure in the Atlantic Heathland trajectory (Table 1, Fig. 1)

period	phase type
1963–1969	high velocity linear phase under steady proportionality
1969–1981	non-linear, turbulent phase, following process occlusion brought on by critically reduced carrying capacity in the site. The process briefly returns to the linear phase around 1976, 1977.

The terminology used in this table is reminiscent of comparing the vegetation process to fluid flow where laminar and turbulent phases alternate (Anand and Orłóci, 2000; Orłóci et al., 2002b). Margalef's (1989) vortex analogy should be mentioned at this point, not conceding that the choice of 'laminar flow' and 'turbulence' is a far more expressive metaphor.

the Atlantic Heathland time series, occurring around time step 7, divides the trajectory into two phases, one linear and the other chaotic (Fig. 1, Table 8). Several major attractor events and many lesser ones are expected in long trajectories. The detection of these may be based on visual inspection of the trajectory mapping in 3d phase space, which is not generally reliable, or on the sharply projecting apices (positive and negative) of the SSD, D , AN , V , A , H , or I graphs. At the resolving power of the Hanging Lake spectrum (Fig. 2), the high peaks (Figs. 3 and 4) have a clear correlation with the dramatic global temperature events.

6.2 Determinism

Process determinism is measurable as the consistency by which the process steadfastly holds a directional time-forward momentum. In the case of a trajectory mapping we look for how well the trajectory keeps to a steady geometric course with reference direction in the preceding time segment. The first reference direction is set by the first time-step X_0 to X_1 . A directional change occurs when compositional change occurs in the paleocommunity that involves changing the established proportionality of the taxa. When proportions change, the trajectory line will break away from the immediate reference direction at some angle. This signals re-aiming by the process to a new attractor, and by the same token, an interruption of the trajectories' time forward momentum. The pattern emerging in one time step may of course be negated in the next.

The question to be put is whether proportionality loss comes by pure chance or by directional change in the forcing conditions. Pure chance is inherent in the Brownian random walk and also in the behaviour of Lorenz's 'strange attractor' (Lorenz, 1963; May, 1976, 1987; May and Oster, 1976; Schroeder, 1991). Should this be behind a directional change, one would be dealing with the case of the fractal (Orłóci, 2001a). If the cause is not pure chance but also involves a deterministic component, one would have the case of stochastics (Orłóci, 2001a). How to tell which one is the case, stochastics or the fractal? Very likely stochastics when we observe a natural trajectory. That is to say, chance variation convoluted with process determinism as dictated by Nature's bio- and physico-ecological laws. All said, a main task of trajectory analysis is the isolation of the two sources of variation in the hope of arriving at an ecologically informative inference. What could cause process determinism? An éclat high-level example is climate warming and cooling which in turn, on the scale of the resolving power of most palynological spectra, are dependent on cyclic regularities in the Earth's motions (Milankovitch, 1941). On that scale of resolution, attractor migration may appear as if a new life zone migrated into the site vacated by an earlier one.

Regarding the interpretation of directional change, the frequency with which it occurs relative to unit time is important. Phase change in the Heathland trajectory from linear to chaotic is triggered by the severe reduction of the site's carrying capacity. In general, the strengthening of the reign of chance over compositional transitions, with nadir in total chaos, is the antithesis for a strengthening of determinism. This is said in general terms, the specific question remains then: how to measure process determinism? The measurement is simplest if the determinism expected has a surrogate, well-defined reference configuration. Among different possibilities for surrogates two stand out by being at the extremes:

(1) Stationary Markov chain model. The Markov model can be made operational by specifying two distance

configurations. One is \mathcal{A} , defined for the points (paleorelevés) of the fitted stationary Markov chain and the other \mathbf{D} defined for the observed trajectory points. Relevant details are presented by Orlóci et al. (1993, 2002a) and an overview in Web Only Appendix B. The actual technique goes like this: (a) Translate the time series to equal time steps (Web Only Appendix H) and fit the Markov chain to the adjusted data (Orlóci et al., 1993); (b) Compute distance configuration \mathcal{A} for all pairs of the fitted discrete Markov trajectory points and D for all pairs of observed trajectory points; (c) Compute the stress coefficient $\sigma(D, \mathcal{A})$ which measures the incompatibility of structures D and \mathcal{A} ; (d) Find α , the probability that stress by chance alone will be equal to or greater than the observed $\sigma(D, \mathcal{A})$. We use a Monte Carlo experiment for this under the assumption of a purely chance-driven compositional transition process. Since a large $\sigma(D, \mathcal{A})$ means low Markovity and a small, α clearly, the right tail probability scale is used which then runs counter to the goodness of fit of the Markov chain. Therefore, the smaller the value of $1-\alpha$ the stronger the assumption of Markovity is.

Numerical results are given in Table 9. These indicate that determinism of the stationary Markov type is a reasonable assumption for the linear phase in the Atlantic Heathland, and for the Hanging Lake trajectory, but not for

the non-linear phase of the Atlantic Heathland and not for the third example, the Lagoa das Patas trajectory (graphs not given) taken at its full length.

(2) Rank-order model. The rank order of the observed distances in \mathbf{D} is compared to the rank order of time differences in \mathcal{A} . The comparison function is the product moment correlation coefficient $r(\mathbf{D}, \mathcal{A})$. Probabilities (α) are determined for $r(\mathbf{D}, \mathcal{A})$ in a Monte Carlo experiment under the assumption of a pure, chance-driven compositional transition process. The right tail α discriminates against the assumption of directedness. That is, the greater the alpha is, the weaker the trajectory's directedness. It can be seen from Tables 9 and 10 that the two methods lead to similar conclusions in the cases examined.

6.3 Periodicity

Given a sufficiently long period of time, the taxa and the trajectory variables are expected to show periodicity other than pure random oscillations. Taxon-level periodicity is illustrated by the deviation graphs of Fig. 3 and sum of squared deviations, distance, velocity, acceleration, angle, entropy, and information periodicity at the trajectory level in Figs. 3 and 4. In all cases the type of periodicity is very irregular. This holds true for amplitude and also for

Table 9 Probing the sample trajectories for the intensities of Markov type determinism

trajectory	time period /yr	stress observe δ	stress expected	variance of expected stress	probability discriminating against Markovity ($1-\alpha$)	strength of Markov type directedness
Atlantic Heathland	1963–1982	0.244	1.0685	1.41896	< 0.001	intense
	1963–1970	0.162	0.8094	1.32603	< 0.001	intense
	1970–1982	1.451	1.4279	1.59013	0.635	weak
Hanging Lake	0–41000 BP	1.788	1.8035	0.73826	< 0.01	strong
Lagoa das Patas	0–42000 BP	1.872	18.69	0.62082	0.650	weak
RND	0–10000*	2.000	2.000	0.0000	1.000	nil

Consult Web Based Appendix B for the Markov chain and Tables 1 and 2 for site information. Probabilities and expected values are determined in Monte Carlo experiments. RND signifies a long series of random numbers, simulating purely chance-driven compositional transitions. For interpretation of α , see the main text. *: Model time.

Table 10 Probing sample trajectories for the strength of directedness based on the rank correlation $r(\mathbf{D}, \mathcal{A})$

trajectory	time period /yr	rank correlation observed $r(\mathbf{D}, \mathcal{A})$	rank correlation expected	variance of expectation	probability α discriminating against rank order type directedness	strength of directedness
Atlantic Heathland	1963–1981	0.5810	0.0005	0.0612	< 0.001	intense
	1963–1970	0.9291	0.0007	0.0377	< 0.001	intense
	1970–1981	0.2938	-0.0069	0.0142	0.537	moderate
Hanging Lake	1–41000BP	0.2462	0.0004	0.0007	< 0.01	strong
Lagoa das Patas	0–42000BP	0.0001	-0.0001	0.00004	0.461	low
RND	0–10000*	-0.0200	0.0001	0.4	0.999	nil

The rank correlation is a product moment applied to the observed distance matrix \mathbf{D} in which distances are replaced by their order and distance matrix \mathcal{A} of time ordinals. Probabilities α and expected values are generated in Monte Carlo experiments with series of random numbers (RND) simulating purely chance driven compositional transitions. Since we intend to determine the strength of the correlation in probability terms and since α is a right-tailed probability, low α values indicate increasing directedness. *: model time.

wavelength of entropy and complexity (Anand and Orłóci, 1996) in general as seen in Figs. 9 and 10. The local peaks and lows are linked to irregularly occurring forcing events. For example, in the Atlantic Heathland case the peaks and lows track the community emerging from heavy grazing in 1964, ending in severe fire; critical process occlusion occurring in 1968, the year after dramatic reduction of bare ground surface; and the draught years 1976–1977. Tables 11 and 12 contain correlation values for H , C , and V in Figs. 9 and 10. In the Atlantic Heathland, the overall relationship is positive, but varies considerably depending on the period. As seen in Table 12, a different picture emerges when we take the relationship of H , C , and V in the long trajectory of Lagoa das Patas. It is interesting to find a negative correlation of V with both H and C dominating the frequency distribution. This point supports the proposition that decreasing velocity brings on increasing diversity and thus increasing process stability.

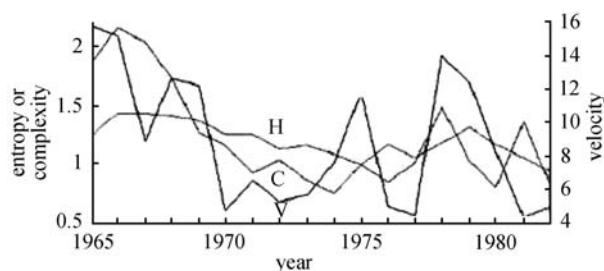


Fig. 9 Evolution of disorder-based compositional diversity H , Anand's structural complexity C , and composition based process velocity V in the Atlantic Heathland site (Table 1). Graphs are scaled for clear viewing. Reference and explanations are found in the text and in Web Based Appendix G.

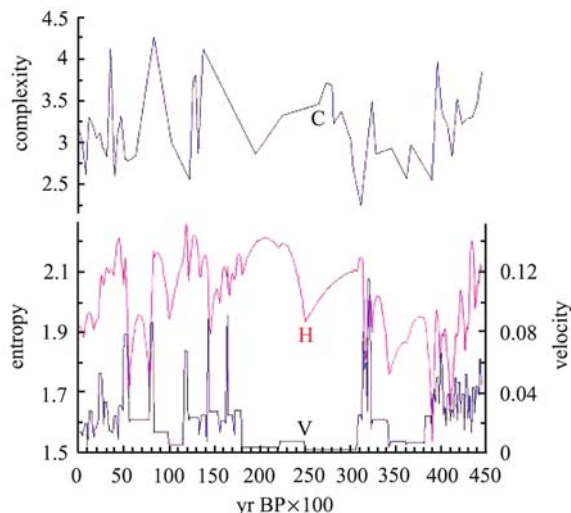


Fig. 10 Long-term evolution of disorder-based compositional diversity H , Anand's structural complexity C , and composition based process velocity V in the Lagoa das Patas site (Table 2). Graphs are scaled for clearer viewing. Explanations are found in the text and in Web Based Appendix G. Equal time step (100 yr) transformation applies (Web Based Appendix H).

6.4 Shape complexity

Any ordered series of numbers will define a unique graph and the graph will have a shape of certain complexity. This is in fact a measurable property. The question is which complexity scalar to choose to complete the task. The choice has to be specific to the properties we wish to emphasize. Obviously, shape and order are related. But we already dealt with order/disorder related issues. So it is tempting now to isolate the property 'complexity' from order and to emphasize aspects of 'shape'.

To specify further what properties of the complexity scalar should have, we set limit values, 1 for graph shape that is maximally simple, for example a straight line, and 2 for graph shape that is maximally complex, that is, beyond which in theory the index could not have values, such as in the case of the path of a particle in true Brownian motion. The Hausdorff (fractal) dimension (Schroeder, 1991) satisfies our requirements. Since there are many ways to implement it, we will follow a technique that requires length measurements (Web Only Appendix E). We refer to Mandelbrot (1967, 1983) for original example of the use of the fractal dimension to fix a number for the complexity of the British coast line.

The Hausdorff dimension D has two defining relations: $L(r) = L_0 r^{1-D}$ and $L(r) \approx r^{1-D}$. Based on these, we can calculate D in a regression analysis in the manner of $D = 1 - b$, where L_0 is the true length of graph, the sum of the trajectory edges' lengths; b is the regression coefficient of $L(r)$ on r ; r is the calliper width setting by which end-to-end graph length is measured.

The measuring convention requires that the measured length be an integer multiple of the calliper width. For example, when we say k in $L(r) = kr$ is an integer, it means that we disregard the last remaining fractional length less than r . The real measurements below illustrate this clearly (Table 13).

The true length of the graph is 9748.5464 in co-ordinate units. Note, the increments in r by factor 1.1 continue until r is smaller than or equal to the largest diameter of the trajectory's point configuration, *i.e.* smaller or equal to the largest value in the trajectory's full dimensional distance matrix. We use a True Basic code for analytical simulation of the calliper's walk through the graph's entire length recursively at increasing r . D indicates the Hausdorff dimension, a number between 1 and 2, inclusive. b indicates the regression coefficient of $\ln L(r)$ on $\ln r$ over the full range of r .

The length based Hausdorff dimensions are listed for seven graphs in Table 14. What can we deduce from the Hausdorff numbers? Considering the Atlantic Heathland trajectory (AHL), the trajectory's total length is 1714.7973 units, of which calliper width at its highest ($r = 117.39085$) recovers 586.95426 units. Since the next increment in r would take one arm of the calliper beyond the largest diameter of the trajectory's distance configuration, we

Table 11 Product moment correlation coefficient (r) of the Atlantic Heathland graphs H , C , and V in Fig. 9

trajectory	interval	$r(H,C)$	$r(H,V)$	$r(C,V)$
Atlantic Heathland	1964–1970	0.6352	0.3477	0.7034
	1970–1978	0.2616	−0.052	0.6676
	1978–1981	0.0034	0.9355	−0.2701
	1964–1981	0.6758	0.5364	0.6351

H : disorder based compositional diversity, Rényi's (1961) entropy of order one; V : composition based process velocity (see the text); C : Anand's structural complexity; Explanations are found in the text and in Web Based Appendix G.

Table 12 Values of the product moment correlation (r) of the H , C , V graphs in Fig. 10 for Lagoa das Pata (Table 2)

	$r(C,V)$	$r(H,V)$
correlation	−0.29	−0.24
frequency of positive correlation values /%	11	15
frequency of negative correlation values /%	86	76
frequency of zeros/%	3	9

The values given are regression estimates at $BS=1$. V : composition based process velocity; H : disorder based compositional diversity; C : Anand's structural complexity. Equal time step (100 yr) transformation applies (see technique in Web Only Appendices F and H).

Table 13 Interpretation of the measuring convention $L(r)$

r	1	1.1	1.21	1.331	...	207.96506
k	9706	8822	8018	7288	...	8
$L(r)=kr$	9706	9704.2	9701.78	9700.328	...	1663.7205

consider the r value 117.39085 final. The number of $L(r)$ values generated is 51 and the corresponding Hausdorff dimension $D = 1.1429857$. After 1000 random rearrangements of the trajectory points and as many new D values, we have the mean Hausdorff dimension of the randomized trajectory $ED = 1.1730802$. The 99% confidence limits for ED are 1.1068629 and 1.2467207 (in the same units). These limits completely enclose $D = 1.1429857$. We conclude from these that the trajectory's temporal order

has no significant effect on its shape complexity. In all other cases, except Lynches Crater, similar conclusions are reached. For that reason we conclude as a generalization that temporal order is not a significant determinant of shape complexity. Order discounted, the responsibility must lie with high evenness and regularity. As a matter of interest, cumulative Hausdorff dimensions are given for the Vostok graph (Fig. 11) in Table 15. The values show remarkable stability.

Table 14 Hausdorff (fractal) dimensions of trajectories

site	true length of the trajectory	range of calliper width r	trajectory length $L(r)$ at upper limit of r	Hausdorff dimension D	simulated mean Hausdorff dimension MD	variance V of MD	upper limit of simulated 99% confidence interval	lower limit of simulated 99% confidence interval
LdP	17048	1–445	2674	1.13	1.13	4.07E-06	1.08	1.20
AHL	1714	1–117	586	1.14	1.17	1.047E-05	1.11	1.25
CMB	44171	1–64	405	1.24	1.17	8.72E-04	1.11	1.27
RS	9672	1–955	1911	1.24	1.24	1.27E-06	1.22	1.27
TdF	8010	1–228	686	1.26	1.23	9.06E-06	1.17	1.32
HL	25064	1–368	368	1.29	1.26	8.98E-06	1.15	1.35
LyC	13656	1–207	207	1.47	1.28	1.48E-05	1.19	1.38

Site descriptions are in Tables 1 and 2. Site symbols: LdP, Lagoa das Pata, RS, Rusaka Swamp; CMB, Cambará; TdF, Alberton, Tierra del Fuego; HL, Hanging Lake; AHL, Atlantic Heathland; LyC, Lynches Crater. MD and V are determined in 1000 iterations, each involving a complete random rearrangement of temporal order in the trajectory and calculation of a new D . Raw data is adjusted to 100 maxima. The calculations used eight digits. Only rounded numbers are given. See the main text for method.

Table 15 Cumulative averaging of the Hausdorff dimension D for Vostok (Fig. 11)

step n	calipper width r	graph length $L(r)$	regression coefficient b	Hausdorff dimension D	cumulative mean of D
1	1	422766			
2	1.1	422757.5	-2.11E-04	1.000211	1.000211
3	1.21	422755.9	-1.26E-04	1.000126	1.000169
4	1.331	422752.2	-1.07E-04	1.000107	1.000148
37	30.91268	416146.5	-3.17E-03	1.003169	1.000744
38	34.00395	414780.2	-3.62E-03	1.003623	1.000822
116	57565.04	402955.3	-5.25E-03	1.005252	1.012602
117	63321.54	379929.3	-5.35E-03	1.005352	1.012539

Total graph length: 422769.2 units. After 117 steps the maximum dimension of the graphs point configuration is reached. Last line contains the final b and D based on the regression of 116 values of $\log L(r)$ on the same number of values of $\log r$.

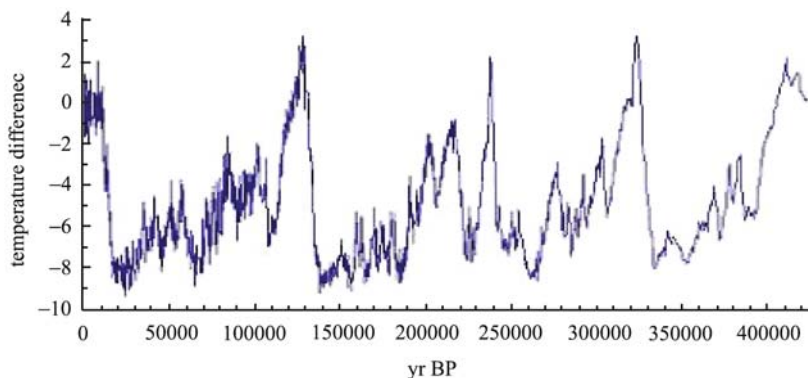


Fig. 11 The Vostok temperature graph after Petit et al. (2001). The Vostok temperature graph spans 3311 readings over 423 thousand years. Relevant statistics are given in Table 15.

6.5 Attractor migration in time and space

Palynological spectra and other vegetation time series capture temporal transitions in the regional vegetation at the core site. Since the core site is but a single pixel in the global vegetation mosaic, the compositional transitions revealed have decreasing relevance as the distance from the core site increases. To improve the perception about transitions in the global pattern's attractor, several core sites have to be examined. One way of doing this has an example in Fig. 1.5 of Delcourt and Delcourt (1987). Their figure, redrawn in our Fig. 12, first row second graph, traces temporal migration of entire vegetation zones in Eastern North America over a 20 kyr period. Based on the Delcourt-Delcourt method one can thus identify the geographic as well as temporal migration of the attractor condition. Upon the examination of transition properties, some interesting facts emerge: The vegetation map in the first row left in Fig. 12 captures the current, regional mosaic of formations in Eastern North America. How did this mosaic come about? Based on the time-by-latitudes graph (2nd row in Fig. 12), the mosaic's evolution is traceable from 20 kyr BP to the present. Regarding the

expansion of the Eastern Cool Temperate Deciduous Forest (D), past and projected transition rates, and lower atmospheric temperature values are juxtaposed. The projected rates are estimates under an assumed, very conservative 3.6°C/100 yr rate of climate warming (Mason, 1990, Intergovernmental Panel on Climate Change, 2001, 2007). We use the Vostok temperature set as proxy for past temperatures after some adjustment. This requires explanation.

The Vostok temperature series covers past thermal events in the Antarctic from as far back as 422 kyr BP. Deuterium based temperatures are given for the thermal inversion level (Petit et al., 2001) expressed as differences from the current inversion level temperature (approximately -82.5°C). To transfer inversion level temperatures to ground temperatures Petit et al. (2001) use a factor, 12/8 = 1.5. In this, 12 is an inversion level temperature amplitude and 8 the ground surface temperature amplitude in celsius degrees. The factor of passing from the Vostok surface temperature (-55°C) to global mean temperature (15°C) is 249°K/179°K = 1.39. We take an even 2 as the factor of combined transformation from Vostok inversion to global mean surface temperature.

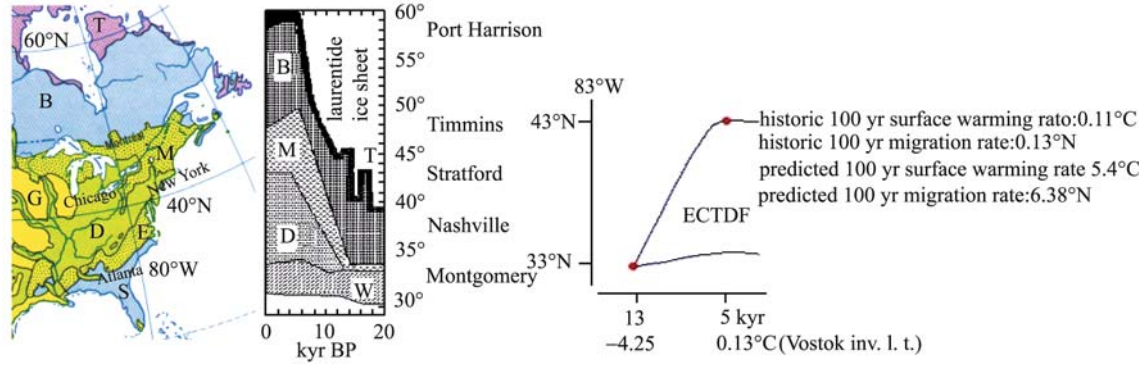


Fig. 12 Vegetation map of Eastern North America (after Espenshade and Morrison, 1990, modified) and graphs showing the historic expansion of the Eastern Cool Temperate Deciduous Forest over latitudes and time. Statistics in graph follow Orłóci (2008). Second graph is adapted from Delcourt and Delcourt (1987, Fig. 1.4) with modifications. The rates shown for temperature are from Vostok transformed to global mean average surface rates and adjusted to show the effect of the local thermal flux rate (Orłóci, 2008). Compare the projected rates to the actual historic rates to see the wide gap, and the enormity of the likely consequences. Formation symbols: T: tundra; B: Boreal Forest; M: Mixed Conifer-Northern Hardwood Forest; D: Cool-Temperate Deciduous Forest; E: Warm-Temperate South-Eastern Evergreen Forest; S: Subtropical. See Table 18 for warming rates and Braun (1950) for detailed description of D.

Table 16 Frequency distribution of 100 yr warming rates based on 4228 steps each 100 yr wide in the complete Vostok set (Fig. 14)

upper class limit for 100yr Vostok inversion level warming rates/°C	upper class limit for 100yr global average surface warming rates/°C	cumulative frequency	frequency	proportion
-2	-4	0	0	0.0000
-1.6	-3.2	1	1	0.0002
-1.2	-2.4	7	6	0.0014
-0.8	-1.6	25	18	0.0043
-0.4	-0.8	212	187	0.0442
0	0	2239	2027	0.4794
0.4	0.8	4001	1762	0.4167
0.8	1.6	4201	200	0.0473
1.2	2.4	4223	22	0.0052
1.6	3.2	4225	2	0.0005
2	4	4228	3	0.0007
2.4	4.8	4228	0	0.0000
			4228	1.0000

Table 17 100 yr warming/cooling rates for Vostok (Fig. 11) within periods as identified

period/TBP yr	warming during period/°C	cooling during period/°C	period length/yr	100 yr inversion level rate/°C	100 yr surface rate/°C
334600–323695	10.89	—	10905	0.0999	0.20
323695–307005	—	9.0900	16690	-0.0834	-0.17
139643–128501	9.92	—	11142	0.0910	0.18
128501–108538	—	9.6300	19963	-0.0883	-0.18
16949–13949	6.26	—	3000	0.2087	0.42
13949–13549	—	1.7400	400	-0.4350	-0.87
12649–4949	3.25	—	7745	0.0420	0.084
12649–11149	4.73	—	1500	0.3153	0.63
11146–10649	—	1.5100	497	-0.3058	-0.612
8619–8135	2.79	—	484	0.5764	1.15

See the text for technique to determine surface rates.

For clarity we give the following explanation of the Vostok Deuterium scale and its relationship to the temperature scale. The scale has basic definition as a ratio $D = 1000 [(R_{\text{observed}} - R_{\text{ocean}})/R_{\text{ocean}}]$, in which the R values accord with the ratio of hydrogen and heavy hydrogen, $R = {}^1\text{H}/{}^2\text{H}$. The regression coefficient of D on temperature in $^{\circ}\text{C}$ is $b = (23 + 170)/(28 + 8) = 5.3$. The regression coefficient is based on the 1984 Gray and Song graph given in Schweingruber (1996, Fig. 20.11). The current Vostok surface temperature is -55°C and the current inversion level temperature is $1.5(-55) = -82.5$. This has a D equivalent of $5.3(-82.5) = 437.3$, a value close to -438 , the very first D entry in the Vostok records (Petit et al., 2001). The temperature values T in the Vostok records are given as approximately $T = D/5.3^{\circ}\text{C} - 82.5^{\circ}\text{C}$.

A closer examination of the Vostok temperature set reveals facts that put the forgoing predicted conditions into perspective. It is interesting to observe the frequency distribution of warming rates (Table 16) over the entire length of the Vostok series, converted to mean global temperatures. About 90% of the time steps sported a 100 yr warming rate between -0.8°C and 0.8°C . Only 5 times did the 100 yr warming rate exceed 3.2°C . In these terms the predicted 100 yr warming rate of 3.6°C is simply colossal and highly unnatural. Equally interesting are the warming rates in Table 17 for selected periods. The highest 100 yr

warming rate found is 1.15°C . Most observed rates are much lower. Compared to these, the projected 100 yr warming rate of 3.6°C —while modest among the projections that float between 1°C and 6°C or higher—is colossal in comparison to most historic rates with a huge potential when compared to the historic consequences of sustained warming at past rates in the fractions of just one Celsius degree.

To complete this section, now the potential effects of predicted global warming is examined on local heat flux at specific locations in the Delcourt and Delcourt (1987) transect belt. The method follows Orlóci (1994) with revisions. The thermal flux rate is the extent to which a one-degree rise in the global mean annual temperature works its way into the measured local temperature. To calculate a site's thermal flux rate, we recognize its vegetation zone, say Z , which had its historic southern limit (lower elevation limit or any target limit) at latitude HSL when the global mean annual temperature was $HGAT$. Suppose that the same zone has its current southern longitudinal limit at CSL where the local mean annual temperature currently is LAT_{CSL} . The current local mean annual temperature at HSL is $CLAT_{HSL}$ and at CSL is $CLAT_{CSL}$. The current global mean annual temperature is $CGAT$. Based on these, the thermal flux rate for Z is:

$$TER = \frac{CLAT_{HSL} - CLAT_{CSL}}{CGAT - HGAT} = \frac{\text{thermal displacement of southern limit or lower elevation limit}}{\text{past global thermal rise}}$$

This expresses the thermal flux rate at the historic southern limit. Note the equation's reliance on the vegetation zone as indicator of thermal flux. Table 18 gives predictions for thermal flux rates (TFR) in the sites, assuming a $3.6^{\circ}\text{C}/100$ yr global sustained warming rate. It is seen that the

predicted thermal flux rate reaches extremely high levels with increasing N latitude. The rates are specific to sites and not transferable between the sites.

When a general circulation model predicts a D amount of rise in the global average temperature, the predicted

Table 18 Local thermal flux rates after Orlóci (1994), modified

vegetation	Arctic Tundra	Boreal Forest	Mixed Conifer- Northern Hardwood Forest	Cool Temperate Deciduous Forest	Warm Temperate Evergreen Forest	Tropical Alpine Tundra (3300 m)	Tropical Subalpine Vegetation (2000 m)
1	climatic station Port Harrison, Qu.	Timmins, Ontario	Stratford, Ontario	Nashville, Ten- nessee	Mobil, Alabama*	Mauna Kea, Hawaii	
2	N.lat.	58°26'	48°31'	43°22'	36°10'	31°42'	19°49'
3	AMP mm	372	711	773	1144	1439	510
4	AMT/ $^{\circ}\text{C}$	-7.5	1.3	8.3	15.6	19.8	0
5	TFR	2.53	1.92	1.5	1.1	0.7**	0.92
6	TR	9.1	6.9	5.5	3.9	3.6	3.32
7	EAMT	1.6	8.2	13.8	19.5	23.4	3.32

Computation of the thermal flux rates (TFR) is explained in the text. Find an original description of the technique in the 1994 paper's pre-publication manuscript, downloadable from the address given (<http://ecoqua.ecologia.ufgrs.br/~lorloci/koa>). Biome boundary shifts in North America are based on the Delcourt and Delcourt graph (1987). The Mauna Kea vegetation records follow Krajina (1963). Data in rows 3 and 4 follow Walter et al. (1975), except in the last two cells which are from Krajina (1963). Abbreviations: AMP, annual mean precipitation; AMT, annual mean temperature; TFR, local thermal flux rate; TR, local temperature rise; EAMT, expected annual mean temperature under the Manabe et al. (1990) scenario of 2.5°C temperature rise in 70 years (equivalent to $3.6^{\circ}\text{C}/100$ yr rate).

*: Not a climatic limit; **: Extrapolation according to $TFR = -1.32071 + 0.065379X = 0.7$. The Mauna Kea site is not included in the calculations of regression constants.

local temperature is $CLAT + TFR \times D$. This value is pointing at specific thermal conditions on the gradient, under which a replacement vegetation model is sought for prediction of a future vegetation state. It should be clear that the success of the replacement vegetation model will depend not only on the new thermal conditions in the site, but also on the local precipitation, soil moisture, and other factors, such as photoperiod conditions, pedogenic conditions, and very much on the time required for the model vegetation to arrive in the site. It is noted further that the predicted conditions may identify model vegetation far outside the floristic region. When this happens, the zonal vegetation collapse under excessive loss of species populations is predicted.

7 Closing remarks

The analytical approach embraced in this paper is constructed on the basis of primary, secondary and scale variables. The primary variables capture the vegetation process in its details, but its fineness depends on the precision of observations. The secondary variables are transformations of the primary variables, and as such describe process structures. The bottom line of precision in this depends on the primary variables' precision. The scale variables control the power and quality of resolution. We refer to the methodology itself as trajectory analysis to emphasize the underlying model from which we infer about the process properties.

Four key phrases capture the essence of trajectory analysis: process description by surrogate trajectory characteristics, reasoning through scales, controlled scale dependence, and open analytical environment. Since the trajectory characteristics are manifestations of the integral process, they apply directly to the integral vegetation process. The problem of re-integration of the elementary processes linked to individual elementary variables is facilitated on any level of detail between total reductionism to total holism. The scale variables allow in a totally controlled manner the manipulation of the resolving power of the analysis and through this the capture of the process in its full complexity. If the early succession studies had access to these devices, much of the frivolous arguments about the climax could have been avoided.

Trajectory analysis can incorporate explanatory variables, either as factor variables that force compositional and functional transitions in the vegetation or as statistical criteria that allow probabilistic tests to be performed on the process related hypotheses. Based on these, the strength of statistical linkage can be established between response and factor variables. It is exactly this aspect of trajectory analysis that revealed novel ecological facts in the examples. One of the most consequential is the discovery that process acceleration/deceleration occurs under climate warming and the exact type of this is formation specific.

Specificity is in the sense that the humid regions suffer significant acceleration of compositional transitions and intense destabilization under sustained climate warming, while the arid regions experience the opposite.

The flexibility of trajectory analysis has to be emphasized. Flexibility makes possible the incorporation of new techniques—novel or imported—under the same conceptual umbrella. Expansion by introduction of new variables as well as new scaling functions, as needed, is facilitated. Incorporation of probabilistic testing is facilitated either as in the Monte Carlo procedures or in the axiomatic distribution based procedures of conventional statistics. To serve ecological users interested in ordered time-static series, trajectory analysis does not have to be changed in other than a semantic context.

Acknowledgements The paper's contents evolved at activities in four sites: the Department of Ecology, Universidade Federal do Rio Grande do Sul, Porto Alegre, Brazil; Department of Biology, Laurentian University, Sudbury, Canada; Department of Botany, The University of Hawaii at Manoa, Honolulu; and the Department of Biology, The University of Western Ontario, London, Ontario. I express gratitude to Professors Valério De Patta Pillar, Madhur Anand, Dieter Mueller-Dombois, and Guillermo Goldstine for facilities and facilitation. My deepest expressions of gratitude are due to Forest Eng. Márta Mihály for advice and sustained, unfailing support. Some tables and figures were adopted from earlier versions published in *Community Ecology* with permission.

References

- Anand M (1994). Pattern, process and mechanism—The fundamentals of scientific inquiry applied to vegetation science. *Coenoses*, 9: 81–92
- Anand M (2000). The fundamentals of vegetation change: complexity rules. *Acta Biotheoretica*, 48:1–14
- Anand M, Orłóci L (1996). Complexity in plant communities: the notion and quantification. *Journal of Theoretical Biology*, 179: 179–186
- Anand M, Orłóci L (1997). Chaotic dynamics in a multispecies community. *Environmental and Ecological Statistics*, 4: 337–344
- Anand M, Orłóci L (2000). On partitioning of an ecological complexity function. *Ecological Modelling*, 132: 51–62
- Anderson P M (1988). Late Quaternary pollen records from the Kobuk and Noatak River drainages, northwestern Alaska. *Quaternary Research*, 29: 263–276
- Anderson R S (1993). A 35,000 year vegetation and climate history from Potato Lake, Mogollon Rim, Arizona. *Quaternary Research*, 40: 351–359
- Bartha S, Czárán T, Scheuring I (1998). Spatio-Temporal scales of non-equilibrium community dynamics in abstract coenostate spaces. *Abstracta Botanica*, 22: 49–66
- Behling H, De Patta Pillar V, Orłóci L, Bauermann S G (2004). Late Quaternary Araucaria forest, grassland (Campos), fire and climate dynamics, studied by high-resolution pollen, charcoal and multivariate analysis of the Camará do Sul core in southern Brazil. *Paleogeography, Paleoclimatology, Paleoecology*, 203: 277–297
- Bonnefille R, Riollet G, Buchet G, Icole M, Lafont R, Arnold M, Jolly D (1995). Glacial/Interglacial record from intertropical Africa, high

- resolution pollen and carbon data at Rusaka, Burundi. *Quaternary Science Reviews*, 14: 917–936
- Braun E L (1950). *Deciduous Forests of Eastern North America*. Toronto: Blakston
- Cajander A K (1909). *Über Waldtypen*. Helsinki: Acta Forestalia Fennica 1
- Çambel A B (1993). *Applied Chaos Theory: a Paradigm for Complexity*. New York: Academic Press
- Clements F E (1916). *Plant Succession: an Analysis of the Development of Vegetation*. Publ. No. 242. Washington: Carnegie Institution
- Colinvaux P A, de Oliveira P E, Moreno J E, Miller M C, Bush M B (1996). A long pollen record from lowland Amazonia: forest and cooling in glacial times. *Science*, 274: 85–88
- Cox D R, Lewis P A W (1968). *The Statistical Analysis of Series of Events*. London: Methuen
- Cwynar L C (1982). A Late-Quaternary vegetation history from Hanging Lake, northern Yukon. *Ecological Monographs*, 52: 1–24
- Czárán T (1998). *Spatiotemporal Models of Population and Community Dynamics*. Population and Community Biology Series. Vol. 21. London: Chapman and Hall
- Delcourt P A, Delcourt H R (1987). *Long-term Forest Dynamics of the Temperate Zone*. New York: Springer-Verlag
- Diggle P J (1981). *Statistical Analysis of Point Patterns*. London: Academic Press
- Edgington E S (1987). *Randomization Tests*. 2nd ed. New York: Marcel Dekker
- Espenshade E B Jr, Morrison J L, eds. (1990). *Good's World Atlas*. 18th ed. New York: Rand McNally, 16
- Fekete G (1985). Terrestrial vegetation succession: theories, models, reality. In: Fekete G, ed. *Basic questions of Coenological Succession*. Budapest: Akademiai Kiadó, 31–63 (In Magyar)
- Fekete G, Virágh K, Aszalós R, Orlóci L (1998). Landscape and ecological differentiation of *Brachypodium pinnatum* grasslands in Hungary. *Coenoses*, 13: 39–53
- Foeli E, Orlóci L (1985). Species dispersion profiles of anthropogenic grasslands in the Italian Pre-Alps. *Vegetatio*, 60: 113–118
- Gleick J (1987). *Chaos: Making a New Science*. New York: Penguin Books
- Goodall D W (1967). Computer simulation of changes in vegetation subject to grazing. *Journal of Indian Botanical Society*, 46: 56–61
- Goodall D W (1972). Building and testing ecosystem models. In: Jeffers N R, ed. *Mathematical Models in Ecology*. Oxford: Blackwell, 173–214
- Greig-Smith P (1952). The use of random and contiguous quadrats in the study of the structure of plant communities. *Annals of Botany*, 16: 293–316
- Greig-Smith P (1957). *Quantitative Plant Ecology*. London: Butterworths
- Greig-Smith P (1983). *Quantitative Plant Ecology*. 3rd ed. London: Blackwell Scientific
- Hammersley J M, Handscom D C (1967). *Monte Carlo Methods*. London: Methuen
- He X S, Orlóci L (1999). Anderson Pond revisited: the late Quaternary vegetation process. *Abstracta Botanica*, 22: 81–93
- Hill M O, Gauch H G (1980). Detrended correspondence analysis: an improved ordination technique. *Vegetatio*, 42: 47–58
- Jacobs B F (1985). A middle Wisconsin pollen record from Hay Lake, Arizona. *Quaternary Research*, 24: 121–130
- Kerner von Marilaun A (1863). *Das Pflanzenleben der Danauländer*. Wagner: Innsbruck
- Kershaw A P (1994). Pleistocene vegetation of the humid tropics of northeastern Queensland, Australia. *Palaeogeography, Palaeoclimatology, and Palaeoecology*, 109: 399–412
- Krajina V J (1963). Biogeoclimatic zones on the Hawaiian Islands. *Newsletter of the Hawaiian Botanical Society*, 7: 93–98
- Kühler A W (1990). Natural vegetation. In: Espenshade E B Jr, Morrison J L, eds. *Good's World Atlas*, 18th ed. New York: Rand McNally, 8–9
- Legendre P, Legendre L (1983). *Numerical Ecology*. Amsterdam: Elsevier
- Lippe E, De Smidt J T, Glen-Lewin D C (1985). Markov models and succession: a test from a heathland in the Netherlands. *Journal of Ecology*, 73: 775–791
- Lorenz E N (1963). Deterministic nonperiodic flow. *Journal of Atmospheric Science*, 20: 130–141
- Lotka A J (1925). *Elements of Physical Biology*. Baltimore: Williams and Wilkins
- Lozhkin A V, Anderson P M, Eisner W R, Ravako L G, Hopkins D M, Brubaker L B, Colinvaux P A, Miller M C (1993). Late Quaternary lacustrine pollen records from southwestern Beringia. *Quaternary Research*, 39: 314–324
- Manabe S, Bryan K, Spelman M J (1990). Transient response of a global ocean-atmosphere model to a doubling of atmospheric carbon dioxide. *Journal of Physical Oceanography*, 20: 722–749
- 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*. San Francisco: Freeman
- Mandelbrot B B (1983). *The fractal geometry of Nature*. San Francisco: Freeman
- Margalef R (1989). On diversity and connectivity as historical expressions of ecosystems. *Coenoses*, 4: 121–126
- Mason J (1990). *The greenhouse effect and global warming*. Information Office, British Coal, C.R.E. Stoke Orchard, Cheltenham, Gloucestershire, U.K. GL52 4RZ
- May R M (1976). Simple mathematical models with very complicated dynamics. *Nature*, 261: 459–467
- May R M (1987). *Chaos and the dynamics of biological populations*. London: Proc. Royal Soc., Series A, 413: 27–44
- May R M, ed. (1981). *Theoretical Ecology*. Oxford: Blackwell
- May R M, Oster G F (1976). Bifurcations and dynamic complexity in simple ecological model. *The American Naturalist*, 110: 573–599
- McIntosh R P (1985). *The Background of Ecology: Concept and Theory*. New York: Cambridge University Press
- Metropolis N (1987). The beginning of the Monte Carlo method. *Los Alamos Science Special Issue*, 125–130. available at URL: <http://library.lanl.gov/cgi-bin/getfile?00326866.pdf>
- Metropolis N, Ulam S (1949). The Monte Carlo method. *Journal of the American Statistical Association*, 44: 335–341
- Mueller-Dombois D (1992). A natural dieback theory, cohort senescence, as an alternative to the decline disease theory. In: Manion P D, Lachance D, eds. *Forest Decline Concepts*. St. Paul, MN: The Am

- Phytopath Soc Press, 26–37
- Orlóci L (1974). On information flow in ordination. *Vegetatio*, 29: 11–16
- Orlóci L (1978). *Multivariate Analysis in Vegetation Research*. The Hague: W. Junk bv
- Orlóci L (1991a). On character-based community analysis: choice, arrangement, comparison. *Coenoses*, 6: 103–107
- Orlóci L (1991b). Entropy and Information. *Ecological Computations Series (ECS)* Vol. 3. The Hague: SPB Academic Publishing bv
- Orlóci L (1991c). CONAPACK: A program for Canonical Analysis of Classification Tables. *Ecological Computations Series*: Vol. 4. The Hague: SPB Academic Publishing
- Orlóci L (1993). The complexities and scenarios of ecosystem analysis. In: Patil G P, Rao C R, eds. *Multivariate Environmental Statistics*. New York: Elsevier Scientific, 423–432
- Orlóci L (1994). Global warming: the process and its anticipated phytoclimatic effects in temperate and cold zone. *Coenoses*, 9: 69–74
- Orlóci L (2001a). Pattern dynamics: an essay concerning principles, techniques, and applications. *Community Ecology*, 2: 1–15
- Orlóci L (2001b). Prospects and expectations: reflections on a science in change. *Community Ecology*, 2: 187–196
- Orlóci L (2006). Diversity partitions in 3-way sorting: functions, Venn diagram mappings, typical additive series, and examples. *Community Ecology*, 7: 253–259
- Orlóci L (2008). Vegetation displacement issues and transition statistics in climate warming cycle. *Community Ecology*, 9: 83–98
- Orlóci L, Anand M, He X S (1993). Markov chain: a realistic model for temporal coenoses? *Biométrie-Praximétrie*, 33: 7–26
- Orlóci L, Anand M, Pillar V D (2002b). Biodiversity analysis: issues, concepts, techniques. *Community Ecology*, 3: 217–236
- Orlóci L, Orlóci M (1988). On recovery, Markov chains and canonical analysis. *Ecology*, 69: 1260–1265
- Orlóci L, Pillar V D, Anand M (2006). Multiscale analysis of palynological records: new possibilities. *Community Ecology*, 7: 53–68
- Orlóci L, Pillar V D, Anand M, Behling H (2002a). Some interesting characteristics of the vegetation process. *Community Ecology*, 3: 125–146
- Petit J R, Jouzel D, Raynaud D, Barkov N I, Barnola J M, Basile I, Bender M, Chappellaz J, Davis J, Delaygue G, Delmotte M, Kotlyakov V M, Legrand M, Lipenkov V, Lorius C, Pepin L, Ritz C, Saltzman E, Stievenard M (2001). Vostok Ice Core Data for 420,000 years, IGBP PAGES/World Data Center for Paleoclimatology Data Contribution Series 2001-076. NOAA/NGDC Paleoclimatology Program, Boulder CO, USA
- Pielou E C (1977). *An Introduction to Mathematical Ecology*. New York: Wiley-Interscience
- Pillar V D, Orlóci L (1996). On randomisation testing in vegetation science: multifactor comparison of relevé groups. *Journal of Vegetation Science*, 7: 587–592
- Podani J (1994). *Multivariate Analysis in Ecology and Systematics*. The Hague: SPB Publishing
- Rényi A (1961). On measures of entropy and information. In: Neyman J, ed. *Proceedings of the 4th Berkeley Symposium on Mathematical Statistics and Probability*. Berkeley: University of California Press, 547–561
- Reyner J N (1971). *An Introduction to Spectral Analysis*. London: Pion Limited
- Ripley B D (1981). *Spatial Statistics*. New York: Wiley & Sons
- Schroeder M (1991). *Fractals, Chaos, Power laws*. New York: Freeman
- Schweingruber F H (1996). *Tree Rings and Environment Dendroecology*. Swiss Federal Institute for Forests Snow and Landscape Research, Bern: Birmensdorf and Paul Haupt
- Shugart H H, ed. (1977). *Time Series and Ecological Processes*. Philadelphia: SIAM
- Singh G, Geissler E A (1985). Late Cainozoic history of vegetation, fire, lake levels and climate at Lake George, New South Wales, Australia. *Philosophical Transactions of the Royal Society of London Series B*, 311: 379–447
- Sukachev V N (1913). *Introduction to the Study of Plant Communities*. Bibl. St. Petersburg: Natur (In Russian)
- Trewartha G T (1990). Climatic regions. In: Espenchade E B Jr, Morrison J L eds. *Rand McNally Good's World Atlas*. 18th ed. New York: Rand McNally, 8–9
- Trewartha G T (2001). *Global Mechanism of UNCCD*, Via del Serafico 107, 00142 Rome, Italy. Web address: www.gm-unccd.org/English/Field/aridity.htm
- Usher M B (1981). Modelling ecological succession with particular reference to Markovian models. *Vegetation*, 46: 11–18
- Usher M B (1992). Statistical models of succession. In: Glenn-Lewin D C, Peet R K, Veblen T T, eds. *Plant Succession: Theory and Prediction*. London: Chapman and Hall, 215–248
- van Hulst R (1992). From population dynamics to community dynamics: modelling succession as a species replacement process. In: Glenn-Lewin D C, Peet R K, Veblen T T, eds. *Plant Succession: Theory and Prediction*. London: Chapman and Hall, 188–214
- van Hulst R (2000). Vegetation dynamics and plant constraints: separating generalities and specific. *Community Ecology*, 1: 5–12
- Volterra V (1926). 'Variazioni e fluttuazioni del numero d'individui in specie d'animali conviventi', *Mem. Acad. Lincei*, Vol. 2, No. VI. 31–113. Reprinted: 409–448 in Chapman, R.N. 1931. *Animal Ecology*. McGraw-Hill, NY
- von Post L (1946). The prospect for pollen analysis in the study of the earth climatic history. *New Phytologist*, 45: 193–217
- Walter H, Harnickell E, Mueller-Dombois D (1975). *Climate Diagram Maps*. New York: Springer-Verlag
- Watts W A, Bradbury J P (1982). Paleoecological studies at Lake Patzcuaro on the west-central Mexican Plateau and at Chalco in the Basin of Mexico. *Quaternary Research*, 17: 56–70
- Watts W A, Hansen B C S, Grimm E C (1992). Camel Lake: A 40,000-yr record of vegetational and forest history from northwest Florida. *Ecology*, 73: 1056–1066
- Wilkins G R, Delcourt P A, Delcourt H R, Harrison F W, Turner M R (1991). Paleocology of central Kentucky since the last glacial maximum. *Quaternary Research*, 36: 224–239

WEB BASED APPENDICES FOR ...

Trajectory analysis: powerful conceptual tool for understanding change

László Orlóci*

Department of Ecology, Universidade Federal do Rio Grande do Sul, Porto Alegre, RS, 91540-000, BRAZIL

Abstract

The model at the basis of trajectory analysis is conceptually simple. When applied to time series vegetation data, the projectile becomes surrogate for vegetation state, the trajectory for the evolving vegetation process, and the properties of the trajectory for the true process characteristics.

Notwithstanding its simplicity, the model is well-defined under natural circumstances and easily adapted to serial vegetation data, irrespective of source. As a major advantage, compared to other models that isolate the elementary processes and probe vegetation dynamics for informative regularities on the elementary levels, the trajectory model allows to probe for regularities on the level of highest process integrity. Theory and a data analytical methodology developed around the trajectory model are outlined, including many numerical examples. A rich list of key references and volumes of supplementary information supplied in the Web Only Appendices, round out the presentation.

Keywords: Attractor migration; Determinism; Fractal dimension; Parallelism; Periodicity; Phase structure.

APPENDIX A. Eigenanalysis

Eigenanalysis is a tool for rigid rotation of the reference axes to positions of mutual linear independence. Eigenanalysis is the mathematical basis of Principal Components Analysis (Hotelling 1933). PCA was introduced to ecologists in 1954 by David W. Goodall. He called the technique 'ordination'. Others followed using the same terminology for this technique and also for all the others that involved the arrangement of objects on axes as points based on their measured resemblances (Greig-Smith 1957).

*Please direct all correspondence to:

Dr. L. Orlóci, 3-575 McGarrell Pl, London, Canada N6G 5L3. E-mail address: lorloci@nwo.ca

Page for downloadable Web Based Appendices: <http://www.vegetationdynamics.com> link *Appendices for Ta*

There are two basic transformations in most ordination techniques: raw data to resemblance values and resemblance values to new co-ordinates. Given p variables and n units, there will be $p(p-1)/2$ or $n(n-1)/2$ pair-wise resemblance values (depending on the mode of analysis). The second transformation takes the resemblance matrix and produces a new description of the sample in terms of t sets of n ordination co-ordinates. Resemblance functions and ordination techniques are described by Orłóci (1978).

Eigenanalysis is a parsimonious linear method of ordination. It is said to be parsimonious because of the independence of the t co-ordinate sets. In a parsimonious ordination t is smaller, usually much smaller, than p .

Eigenanalysis can lead the same t set of n co-ordinates via two alternative pathways (Orłóci 1966, Orłóci 1967) made obvious by the characteristics equations $\mathbf{AA}'\boldsymbol{\alpha}=\boldsymbol{\alpha}\boldsymbol{\lambda}$ (1) and $(\mathbf{A}'\mathbf{A})(\mathbf{A}'\boldsymbol{\alpha})=(\mathbf{A}'\boldsymbol{\alpha})\boldsymbol{\lambda}$ (2). In these, Matrix \mathbf{A} is a row-centred $p \times n$ data matrix and $\boldsymbol{\alpha}$ is a $p \times t$ matrix of component (transformation) coefficients. The t columns of matrix $\boldsymbol{\alpha}$ are the normalised Eigenvectors of \mathbf{AA}' . Vector $\boldsymbol{\lambda}$ is t -valued with Eigenvalues being the elements. Note: \mathbf{A} has the same definition in (1) as it has in (2). Also, \mathbf{AA}' is a $p \times p$ product matrix defined for the p rows in \mathbf{A} . Further, $\mathbf{A}'\mathbf{A}$ is an $n \times n$ product matrix defined for the n columns of \mathbf{A} . Since the component scores of the n column entities of \mathbf{A} are defined by $\mathbf{Y}=\mathbf{A}'\boldsymbol{\alpha}$, the same scores can be obtained directly by way of (2) as elements in the Eigenvectors of $\mathbf{A}'\mathbf{A}$. Adjustment is required such that the sum of squares of elements in any Eigenvector of $\mathbf{A}'\mathbf{A}$ will be equal to the corresponding Eigenvalue. Analysis via solution (2) is practical when p is larger than n . The following definitions should clarify the technical terminology consistent with our examples:

$\boldsymbol{\lambda} = \lambda_1 \lambda_2 \dots \lambda_t$ – vector of t Eigenvalues

Y_{ij} - component score of relevé j on component i

$\lambda_i = \mathbf{Y}_i \mathbf{Y}_i'$ for any i

$A_{hj} = \frac{X_{hj} - \bar{X}_h}{\sqrt{n-1}}$ – centring of taxon (variable) vector h

X_{hj} – value of taxon h in relevé j

\bar{X}_h – sample mean of taxon h

α_{hi} – component coefficient (direction cosine) of taxon h and Eigenvector i

$\boldsymbol{\alpha}'_i \boldsymbol{\alpha}_i = 1$ for any i

$\alpha'_h \alpha_i = 0$ for any i, h ; this does not exclude the possibility of higher order correlation between components i, h .

$L_{\%} = 100 - E_{\%}$ where $E_{\%} = \frac{100 \sum_{i=1}^{t-k} \lambda_i}{\sum_{i=1}^t \lambda_i}$ – information loss by discarding the last k sets of component

scores, sets arrange in order of decreasing Eigenvalues.

$E_i\% = \frac{100 \lambda_i}{\sum_{i=1}^t \lambda_i}$ – efficiency of i th Eigenvector in accounting for variation in the sample.

The decision to discard the last k sets of component scores is ideally the outcome of a test of the null hypothesis $H_0: E(\lambda_{\tau, k+t}) = \dots = E(\lambda_{\tau})$. Symbol E signifies 'expectation'. When H_0 is true, that is when the k smallest population variances are equal, the point cluster in the subspace of the k components is hyperspherical. Such a cluster shape indicates random variation. Provided that n is large, H_0 is true, the population distribution is multivariate normal, and the sample is taken at random, the test criterion given

by $\chi^2 = -(n-1) \sum_{i=t-k+1}^t \ln \lambda_i + (n-1)k \ln \frac{\sum_{i=t-k+1}^t \lambda_i}{k}$ should have the theoretical Chi-squared distribution (see

Morrison 1976, pp. 296) with degrees of freedom $\nu = \frac{k(k+1)-2}{2}$. The popularity of this technique in

statistical applications justifies reminding users that the regularity conditions as stated may never be fully satisfied in ecological samples, and the test as described may never be fully justified. APPENDICES I and K further clarifies this point.

APPENDIX B. The Markov chain

We mention Feller's (1957) book as the premier reference on Markov chains. We recommend Feller to statistically minded ecologists. The first point to be made to demystify the Markov chain is that any series of numbers (trajectory) generated by the simple recursive law $\mathbf{X}_{t+1} = \mathbf{X}_t \mathbf{P}$ defines a Markov chain. The chain is either "stationary" or "moving". This depends on \mathbf{P} staying constant or changing. No time series will be stationary in Nature for any length of time. The second point is that any square matrix of normalised rows, i.e. the elements add up to 1 in any row, qualifies as \mathbf{P} . Further regarding the symbols, \mathbf{X}_t

is an s -valued data vector. We use it to describe the composition of the paleocommunity at time t . We call such a vector a paleorelevé. We use symbol s to designate the number of taxa, \mathbf{P} is an $s \times s$ transition probability matrix such that an entry in the intersection of row b and column i is the probability of taxon b being replaced by itself ($b=i$) or by another taxon ($b \neq i$) in the next time step. Accordingly, the sum of the elements within any row of \mathbf{P} is equal exactly one. The sum of the elements within the columns may be different.

For purposes of illustration we should consider a simple case where a single taxon is involved: $\mathbf{X}_0 = [10]$ and $\mathbf{P} = [0.5]$. With these given, the stationary Markov chain unfolds like this: 10, 5, 2.5, ... Carrying on in the same manner, we have $\mathbf{X}_{10} = 0.01953125$. The series converges to a value that will not change significantly, no matter how many new steps are calculated. In practice using calculations, the number of steps after which the stable value occurs will depend on the precision of the computer.

The basic problem in Markov chain applications is the determination of the transition probabilities. These cannot be foretold from first principles with any accuracy in natural communities. Therefore they have to be estimated from observations as postdictions. Orlóci et al. (1993) offered a practical solution to the problem of estimation (the first as far as we know) in the context of survey-type, time series vegetation data.

It should be remembered that the Markov chain requires two things to be completely defined: \mathbf{X}_0 and \mathbf{P} . Regarding the frequently asked question whether or not the stationary Markov chain has memory, the answer is definitely yes for path. This is easily seen from the possibility of backtracking, namely $\mathbf{X}_{t,t} = \mathbf{X}_t \mathbf{P}^t$. A word of warning: this will only work if \mathbf{X} has not yet reached stability and definitely not a likely possibility when the natural process is considered.

In practice, the Markov chain is best regarded as a 'moving' chain, *i.e.* a chain with \mathbf{P} not being constant. We should direct attention Orlóci et al. (2002a), from which this short section is adopted.

APPENDIX C. Important data properties

Trajectory analysis takes time series data as input. The main data sources include palynological spectra and permanent plot records like in examples of the main text, and time static surrogates involving space for time substitutions (Kerner 1863, Wildi and Schütz 2000). The specifics of time series actually used in this essay are listed in Table 2 of the main text or mentioned at points of first use in the paper.

The objective being estimation of reality, the outcome will be affected by the way the data are obtained and by the type of the objects actually measured. Errors should be expected. Errors that have to do with measurements and sampling, excluding the barefaced mistakes scientists may commit but cannot explain, are considered further:

1. Measurement errors. These depend on the calibration of the measuring scale. The calibration of a diameter tape, for example, will define the limiting error properties in diameter measurements. Measurement errors have a life history of their own in data analysis. The worst case scenario expects complete error accumulation. Orlóci and Orlóci (1995) discuss error handling in connection with ecological examples and Shchigolev (1960) presents the mathematical theory. A special case of great importance in palynological studies is the accuracy of time measurements. The time scale in all cases except direct observations is the product of some dating technique. Considering the dating problem, the earliest attempt for solution is the space for time substitution technique that Kerner (1863) used. It is still applied in community ecology (Wildi and Schütz 2000, Kovács-Láng et al. 1998). With this the chances for inaccuracy are considerable. Two of its reasons are readily seen: (1) the temporal order established for plant communities on the ground need not be the same as their arrangement in time; (2) the set of community types currently existing in the site need not have all existed in historic times. Dating in palynology has an equivalent to the Kerner method in the disguise of depth for time substitution (Anderson 1896, Lagerheim 1902, von Post 1916, 1946). The invention of the isotope dating techniques (Libby 1955) helped to put dating on firmer ground. But isotope dating is very expensive and has a considerable error margin.
2. Sampling error. This is a direct consequence of unit selection. A typical case is the selection of the site for a sediment core, selection of a horizon in a sediment core, and then selection of the sediment sample within the horizons. The type of choice, random or preferential (see Orlóci and Orlóci 1995, Web Based Appendix J), will determine the properties of the sampling error. When the choice is random, error limits can be set onto the estimates by the usual statistical tools. Error handling becomes a much more difficult problem when the choice is preferential.
3. Taxon identification. There are error sources without direct connection to unit selection or measurement. Taxon identification is an example. Taxonomic expertise is assumed, so the sole remaining source for error is the incompleteness of the specimens available for identification. This problem is epidemic in palynology. The incompleteness of the materials tends to push the systematic level of identification up. The result is heterogeneity in the definition of taxa, clearly seen in Fig. 2 of the main text.

The consequences of the taxonomic error are rather serious in the precision of vegetation description. Clearly, taxon identification is scaling, but in this case the meter stick is the taxonomic system defined along phylogenetic, morphological, or functional lines. To enhance the automation in identification, it is practical to think of a taxonomic system as a character hierarchy in which the character states are arranged in a nested scheme (Orlóci 1991a and Pillar and Orlóci 1993a,b). In this scheme, plant identification reduces to a mapping problem in which the taxa represent distinct runs through the nodes vertically in the hierarchical character scheme. When the number of characters is s and character i has n_i states, the number of distinct runs (the taxa) is $n_1 n_2 \dots n_s$. In this context taxa are character set types (CSTs). When the characters are describing plant functions the taxa represent plant functional types (PFTs, Pillar 1999, Pillar and Sosinski 2003). When the characters emphasise common inheritance, the resulting plant taxa will in essence will qualify as species. The CSTs or PFTs will lump or fragment species, but will not qualify as species. An example for lumping species on a plant functional basis has typical example in the functional type that includes *Cercidium microphyllum* (Leguminosae, Sonoran Desert, North America) and *Bulnesia retama* (Zygophyllaceae, the Monte, Argentina). These species converged into identical functional types through evolution independently in distant geographic sites.

4. Incompleteness of the taxon list. This can be problematic, since the list is the taxonomic basis of trajectory mapping. Topics relevant to sampling and to testing the sufficiency of the taxon list are discussed elsewhere (Orlóci 1978, Orlóci and Pillar 1989, Pillar and Orlóci 1993a,b).

Whatever the error sources, the data is the sole basis to lay open the vegetation process for interpretations in a thoroughly scale-dependent manner. The process properties involved include type and intensity of process determinism; level of random oscillations; phase types and sharpness; complexity level; type and intensity of periodicity; level of parallelism; and attractor migration pattern in time space; identity of forcing factors; hot spots of change in time and space; vegetation indicators of change; mechanisms of governance; and implications in the broader economic and political context.

APPENDIX D. Scale effect on perception

Trajectory analysis leads to inferences about patterns and structures that are not obvious on simple inspection of the data set. But what is also important, the inferences are optimised with respect to the scale effect. To illustrate how the scale effect can manifests itself, it is sufficient just to think of what happens when time step width and period length are changed. The effect is far reaching on the conceptualisation of

the vegetation process with it on the spanning of contradictory theories, while the object medium remains the same:

1. Spatial scale at the site or stand level (forest gap, experimental plot, community patch), time steps in years, period length in decades. This is the time scale used in Table 1 of the main text. What makes the narrow time step so interesting is the increased power of resolution for seeing details. Taking time steps at the year or decade level, the observations will bring aspects of the vegetation process into view whose leading characteristic is dominance sorting. This is also the scale used in early investigations of gap dynamics (Morozov 1912) and more recent forest growth models (Shugart 1984). At this scale vegetation dynamics is readily linked to population processes. The later is highlighted in works with tolerance and inhibition (Connell and Slatyer 1977), life history types (Grime 1977), cohort senescence (Mueller-Dombois 1992), reproductive strategies (Harper 1977), propagule bank composition (Egler 1954). At this scale, the immensity of chance effects cannot be missed.
2. Spatial scales at the landscape or life zone level, period length at the longevity of the same type of climate, time step in decades or centuries. This is the Kernerian facilitation time scale. The period length is already too long for direct observation of the process, and any inference has to rely on palynological fixing or other long-term time series data. Feedback is functioning process, known well to foresters having had to deal with secondary pedogenesis in the form of ortstein development. Typically in this, raw humus accumulates under Pine monocultures planted on sand, the percolating rain water becomes acidic and the leaching of soluble soil components (much clay minerals and sesquioxides) from the upper soil horizon intensifies. The leachates accumulate in a lower horizon and eventually an impervious hard pan, the ortstein layer develops. Soil depth reduced and moisture drainage obstructed, the altered conditions facilitate invasion of species from heath-land communities.
3. Spatial scale global, period length spanning entire cooling and warming cycles, time step in centuries of millennia. This is the level the palynological spectrum captures the vegetation process. For the vegetation process at this level in dynamics the technical literature has two terms: secular succession and precession dynamics. The term 'secular' emphasizes the broadness of the scale while the term 'precession' stresses the scale's coincidence with the climate-defining Milankovitch (1941) cycle in the Earth's motions, *i.e.* the precession of the equinoxes. Delcourt and Delcourt (1987) study of the Late Quaternary vegetation history of Eastern North America is on this level. On this scale, the determinism appears dominant compared to much weaker random oscillations.

The appearance of process properties being unique to the observer's scale(s) do not change the basic fact that the vegetation process viewed is the same unified process on any scale. Only what is being perceived differs. Therefore, an accurate picture of the integral process can emerge only when data and data analysis span the scales at breadth.

APPENDIX E. Shape complexity: the Hausdorff (fractal) dimension

We can characterise a graph by its shape complexity. The question is how to measure shape? Mandelbrot's (1967, 1977) fractal geometry supplies one of the tools. The reader is referred to Schroeder (1991) for a broad perspective on fractals, and to Palmer (1988, 1992), Kenkel and Walker (1993), Scheuring (1993), and Walker and Kenkel (1998) for ecological applications. The following method description is based on an application by Orlóci et al. (2002a). Specifically, we measure complexity as a length related fractal dimension D . We use the power law to link graph length $L(r)$ to scale unit r , and to D in the manner of $L(r) \sim r^{1-D}$ or $\log L(r) \sim (1-D) \log r$ which we may re-write in the form of the regression equation $y = a + b \log r$. Note that $b = 1-D$, the regression coefficient, and $D = 1-b$. Therefore to find the value of D , we have to find the value of regression coefficient b . To this end, we do many length measurements $L(r)$ using different calliper widths r , then regress $L(r)$ on r to find an estimate for $D = 1-b$. The signs of b have to be carefully observed. It will be negative when $L(r)$ increases with r decreasing. This is the case if graph shape is not smooth. D has theoretical values ranging from 1 to 2 depending on the level of complexity. When the graph is the map of a completely random process, the theoretical fractal dimension comes to its maximum 2. The following example illustrates method:

1) Step through the whole length of a graph with different calliper settings to obtain length records. When implementation is analytically based on co-ordinate data (algorithm available from us), the physical dimension of the graph may take any value. Physical implementation can handle graphs only in two dimensional projections. The data set generated in one case is as follows:

Calliper setting (r) mm	2	6	10	16	24	32	40	48
Graph length in number of calliper steps	163	47	24	13	5	3	3	2
Approximate length $L(r)$ mm*	326	282	240	208	120	96	120	96
$\log r$ bits	1	2.6	3.3	4	4.6	5	5.3	5.6

$\log L(r)$ bits 8.4 8.1 7.9 7.7 6.9 6.6 6.9 6.6

- 2) Perform linear regression analysis of $\log L(r)$ on $\log r$ to obtain equation $y=9.087-0.433r$ (see Fig. E1).
- 3) Calculate fractal dimension in the manner of $D=1-(-0.433) = 1.433$

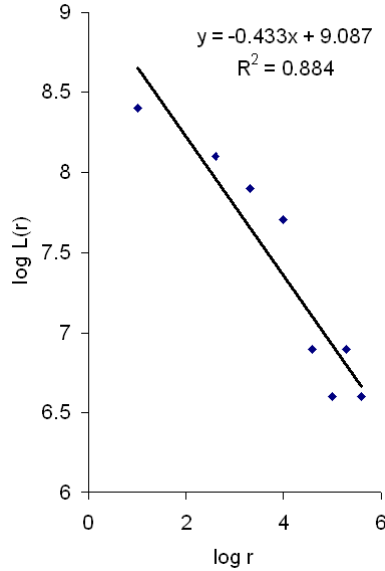


Fig. E1. Regression line fitted to $\log L(r)$ as function of $\log r$ in the data set above. The regression coefficient b (slope of line) is -0.433 , therefore $D=1.433$.

APPENDIX F. Trajectory parallelism

The data set to be used in the numerical example is taken from Fig. F1. The trajectories include clear cases of time-forward progression (A and B), a case in reverse (C), a set of random numbers (D), and two observed cases (E and F). The comparison scalar for any trajectory pair is the topological similarity

coefficient, $TC = \frac{M}{p(s-1)}$. The term '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 not unlike in land surveys, landscape ecology, 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.

The values of TC range from 0 to 1. Values indicate concordance above 0.5 and discordance below 0.5.

The symbols: M - number of matching directional scores on the phase space axes (+, -, 0); 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. The theoretical expectation of TC is taken to be 0.5 under an equal probability law.

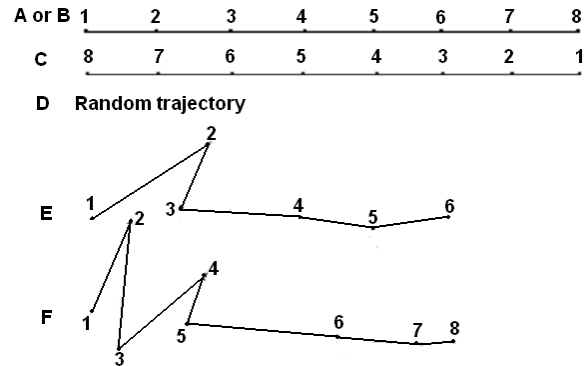


Fig. F1. Sample trajectory mappings (**A,B,C,D,E,F**) based on the data in Table E1. Each trajectory is described by two co-ordinate sets (a,a) . **A** and **B** are identical trajectories, **C** is ordered in the reverse, and **D** simulates chance transitions. **E** and **F** represent two observed cases. The example is adopted from L. Orlóci in Fekete (1998, He and Orlóci 1998) with revisions.

APPENDIX G. Logarithmic measures of diversity and divergence

Rényi's logarithmic expressions

Rényi's (1961) expressions generalize the entropy (H_α) and information (I_α) of distributions in order terms.

The basic definitions are $H_\alpha = \frac{1}{1-\alpha} \ln \sum_{i=1}^s p_i^\alpha$ for entropy and $I_\alpha = \frac{1}{\alpha-1} \sum_{i=1}^s \frac{p_i^\alpha}{\alpha-1}$ for information. Symbols

p_j and q_j are elements in two s -valued distributions, **P** and **Q**. These are identically ordered and have

identical totals, $\sum_{i=1}^s p_i = \sum_{i=1}^s q_i$. The p and q terms are defined as: $p_i = \frac{f_i}{T}$, $T = \sum_{i=1}^s f_i$ and q an expectation

of p . H_α and I_α describe curves as a function of α (Orlóci 1991b). The curves are descending (if not a straight line) for H_α and ascending for I_α . Both are discontinuous at $\alpha=1$, but otherwise continuous over the range from $\alpha=0$ and up. To determine the value of entropy or information of order one, α may be set to a value close to one, say 0.9999 ... or the alternative expressions would have to be used,

$H = \sum_{i=1}^s p_i \log p_i$ and $I = \sum_{i=1}^s \frac{p_i}{q_i}$. What is the significance of α in these expressions? Alpha is the scale

variable and as such it defines an infinite number of possible entropy and information measures. Three of these had special significance for ecologists. Entropy of order zero is the upper limit. Entropy of order one is the Shannon entropy, and information of order one is one half of Kulback's (1959) minimum discrimination information statistic (Kullback et al. 1962). Entropy of order two is a log Simpson index (Simpson 1949). Alpha is useful also in other respects, such as in the detection of the region on the curve at which the diversity or information value becomes stable. This is an advantage when comparisons are made between cases.

H_α has some interesting properties. It has maximum value equal to $\ln s$, corresponding to maximum disorder in the distribution, *i.e.* where $p_1 = p_2 = \dots = \frac{1}{s}$. The degree to which an equidistribution is approached in \mathbf{P} is measured by $E = \frac{H}{\ln s}$ called the evenness or flatness of \mathbf{P} . H_α has minimum value when \mathbf{P} is most contagious, *i.e.*, when $s-1$ of the elements are equal to $\frac{1}{T}$ and the remaining single element is equal to $\frac{T-s+1}{T}$.

I_α is a measure of the information divergence of distribution \mathbf{P} relative to \mathbf{Q} . I_α has minimum value at zero when \mathbf{P} and \mathbf{Q} have element-by-element identity, and maximum value when both \mathbf{P} and \mathbf{Q} are most contagious with the $\frac{T-s+1}{T}$ quantity placed in offset positions. Kullback (1959) discusses regularity conditions under which $2I$ (of order one) is distributed as a Chi-squared variate with $s-1$ degrees of freedom. This property has been used to facilitate statistical tests of hypotheses about the relationship of \mathbf{P} and \mathbf{Q} . As ecological practice has it now, Monte Carlo experiments provide a flexible alternative to finding probabilities for the test. We make reference in this regard to Edgington (1987) for the underlying theory randomization testing in general, to Pillar and Orlóci (1996) and McArdle and Anderson (2001) for contrasting ecological applications, Orlóci (2001a) for consideration of an important dichotomy in applications of Monte Carlo testing with theoretical consequences, and to Orlóci (2006) for useful partitions and examples.

Anand's structural complexity

The complexity measure Δ in $L = \Delta + H$, is of particular interest, since Δ is a complement of disorder based

entropy H . As such it has to do with order, which is structure. The Anand and Orlóci (1996, their Fig. 1) should be consulted on specific details regarding Δ and Anand and Orlóci (2000) on its additive partitioning.

Simpson's index (Simpson 1949)

This is one of the probability-based indices in the manner of $SI = \sum_{i=1}^s p_i^2$, such that $\sum_{i=1}^s p_i = 1$. As given, the index SI expresses the probability of finding a compositional duplicate of the community under the assumption that chance rules composition. When Simpson's function expressed in the manner of $H = -\ln SI$ is equivalent to Rényi's entropy of order 2. Considering that SI at given s has minimum value when $p_1 = p_2 = \dots = \frac{1}{s}$ where H reaches its maximum, SI is a measure of order in the distribution. Energy-focussed ecologists, like Fosberg (1965), have used negentropy in characterisations of SI . 'Negentropy' is a term borrowed from thermodynamics (see Prigogine 1968) where it refers to the available energy in a system. 'Negentropy' is mirrored by 'entropy', the energy that has been spent. In these terms, SI has maximum value when 'negentropy' is maximal, *i.e.*, when the distribution \mathbf{P} is most contagious. The limits are reversed when the Simpson index is inverted in the manner of $SD = \frac{1}{SI}$, which is some measure of 'entropy'. The maximum value of SD is s . As a possible point of interest to some ecologists, we computed values for Simpson's index for the following distributions:

Observed	$\mathbf{P} = [13/16 \ 1/16 \ 2/16]$
Most dispersed	$\mathbf{P}_U = [1/3 \ 1/3 \ 1/3]$
Least dispersed	$\mathbf{P}_L = [14/16 \ 1/16 \ 1/16]$

\mathbf{P}_U and \mathbf{P}_L define \mathbf{P} 's hypothetical upper and lower bounds. The numerical results are presented in Table A8. I leave the interpretation to the interested reader.

Table A8. Values of the Simpson index computed for distribution \mathbf{P} , \mathbf{P}_U , \mathbf{P}_L from above. See earlier sections for the explanation of symbols.

\mathbf{P} \mathbf{P}_U \mathbf{P}_L

$SI = \sum_{i=1}^s p_i^2$	0.50617284	0.33333333	0.62962963
$SD = \frac{1}{SI}$	1.9756098	3.0000000	1.5882353
$SDE = \frac{SD}{s}$	0.65853659	1.0000000	0.52941177
$H_2 = -\ln SI$	0.68087709	1.0986123	0.46262352
$E = \frac{H_2}{\ln s}$	0.61976103	1.0000000	0.42109808

The metric connection

A typical example in the use of the Euclidean metric for diversity measurement is the McIntosh diversity index (McIntosh 1967). The Simpson index is also this kind when given in the form of \sqrt{SI} . Another information theoretical measure, Rajski's (1961) metric is not Euclidean (Orlóci 1978). Additive partitioning can be performed directly on Euclidean metrics in the manner of an analysis of variance, or such as in the method of sums of squares partitioning used in Edwards and Cavalli-Sforza (1965, Pillar and Orlóci 1996, Legendre and Anderson 1999, and McArdle and Anderson 2001). The numbers may not add up in the case of non-Euclidean metrics. To broaden the field covered in this appendix we refer for additional materials in Peet (1974), Pielou (1975), Patil and Taillie (1979), Rao (1982), Juhász-Nagy and Podani (1983), Magurran (1988), Orlóci (1978,1991a), Tóth-Mérés (1997,1998a,b) and Levin (2001).

Table G1. Co-ordinate pairs (a_1, a_2) and directional scores (b_1, b_2) corresponding to the trajectories drawn in Fig. F1. Legend: pa – past; pr – present.

A or B				C				D				E				F			
a_1	b_1	a_2	b_2	a_1	b_1	a_2	b_2	a_1	b_1	a_2	b_2	a_1	b_1	a_2	b_2	a_1	b_1	a_2	b_2
1		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	-
8	+	8	+	1	-	1	-	7	+	6	-	91	+	26	+	120	+	10	+

p_r																			
-------	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--

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

- a. 'Present' is the same time point or the same narrow time interval in all trajectories. When working with paleopollen samples, the top sediment horizon may miss being from the 'present', owing to possible erosion, decomposition, or other processes active in the sediment profile. Because of this, it makes sense to repeat the comparison of the trajectories at varying lag forward or backward. The aim is to maximise TC .
- b. In the interest of finding comparable points on different trajectory lines, we should assume uniform time step widths. See Appendix H for method.
- c. Noting the above, any given state of the process, symbolically X , is endowed with the potent property of being at least to some extent the outcome of random effects, and as such, of being at a locus in phase space unpredictable before the fact. It is justified, therefore, to regard any X as one point in a swarm, any member of which could materialise as an observation, should 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 limits of a tolerance sphere around X in a stepwise manner. 'Tolerance sphere' should not be confused with the probabilistic confidence sphere, belt or interval of statistical estimation. While confidence spheres, belts or intervals help to control the effect of the sampling error, tolerance spheres, belts or intervals help to reduce the effects of small-scale stochasticity in the analysis when making comparisons between the trajectory points. The tolerance sphere is enlarged, starting with radius 0, in small steps up to some arbitrary upper percentage limit defined for the co-ordinate values. All co-ordinates on all axes undergo tolerance transformations with the following possible results for two co-ordinates of the same trajectory on one axis:

0 assigned when the tolerance belt around X includes the chronologically next point on the same axis.

If it does not, then:

+ assigned when the co-ordinate value of the first point given is greater than the co-ordinate value of the chronologically next point (forward).

– assigned when 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 of axes 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-empt the so-called Procrustes manipulations of Shönemann and Carrol (1970). In the present example, the choices are based on local maxima within the 2 x 2 sub-matrices of between-axes correlations as in the following --

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

A quick example will clarify the arithmetic in the calculation of TC . Consider trajectories E and F . Quantity M 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 topologies. The entire set of similarity values and probabilities (α) for the trajectories portrayed above are as follows:

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

The TC values are measures of the strength of topological similarity. The corresponding probability values are generated under the assumption of chance compositional transitions. A TC value is considered

significant if the probability associated with it is small. One more thing is mentioned. The number of reference axes may be unequal in the two trajectories. To resolve the comparison problem, the comparison is limited to axes up to the common number.

e. Order transformation of co-ordinates is implicit when we use +, -, 0 scores. Speaking about process trajectories in real time, trajectory **A** above has a perfect relationship of process time and compositional change in trajectory **B**. Nothing of the sort exists in the case of the trajectory pair **A** and **C**. The value 0.5 is random expectation from first principles. The expectation generated in Monte Carlo experiments may differ. The difference of the theoretical expectation and the experimental expectation is a measure of bias (see table above).

f. 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 for related statistics. Fig. G2 displays the resulting graphs for the trajectories of in Fig. F1.

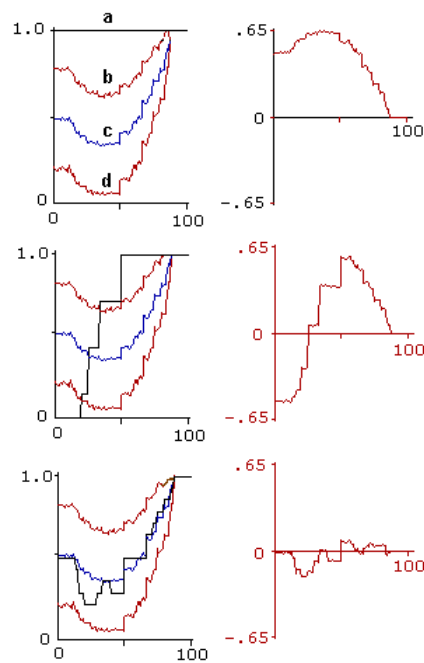


Fig. G2. Graphs of the topological similarity coefficient (*TC*) in left column: *a* -observed value; *c* - random expectation (iterated), *b,d* - limits of the 0.95 confidence belt. Graphs of deviations (*a-c*) are in the right column. On all graphs, the vertical scale is *TC*. The horizontal scale is the tolerance radius from 0% to 100%. Three cases of parallelism are portrayed: (1) Perfect (top graph, trajectories **A** and **B**) indicated by the horizontal *TC* line across the top at 1.0 outside the tolerance radius *b,d*. (2) Complete directional

reversion (middle graph, trajectories **A** and **C**); much negative deviation from expectation. (3) No relationship (bottom graph, trajectories **A** and **D**). The *TC* graph is hugging expectation.

The more adventurous may try an alternative method. In this, after equal time step transformation of the trajectories inner angles are computed in the manner of the angle scalar in the main text. For angles less than 90 degrees minus sign is assigned to points B. Otherwise plus sign is allotted. In other respect the comparison is following the original method.

APPENDIX H. Equal time step transformation

Equal time steps are established by interpolation techniques. The method used here is simplest to explain by example:

(i) Take a set of observations like these:

Depth interval cm	0-10	10-20	20-30	Total
Time interval	5	3	7	15
Pollen count of taxon A	15	6	42	63
Average count per unit time	3	2	6	--

(ii) Arrange the averages into a 15-valued string in the manner of

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

(iii) Selecting 3 basic steps as the common step width, we obtain the interpolated pollen counts

Time step width	3	3	3	3	3	15
Pollen count for taxon A	9	8	10	18	18	63

Clearly, interpolation is smoothing the series and as a possible payoff it clarifies the trend.

APPENDIX I. Non-linearity, prediction, testing

We included materials in this section on non-linearity, on prediction by analogues, and on tests of generality after Orlóci (1993). The discussion of these is not an organic part of the note's main text, but they are target of frequently asked questions.

11. Non-linearity

The point should not be missed that in performance terms the expected type of response individually or jointly is as a rule non-linear under natural conditions. This has consequences as a revelation about the system and about the choice of methods to analyse the system. An idealised two-species case is shown in Fig. I1.

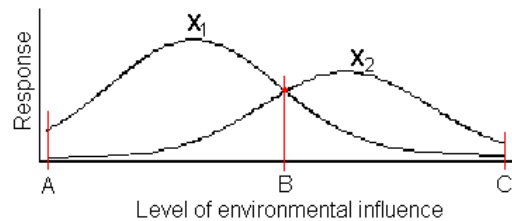


Fig. I1. Idealised species response graphs of the Groenewoud-Whittaker type shown as trend lines without showing stochastic oscillations.

Observing Fig. I1 it should not be missed that under communal existence and natural conditions responses are stochastic, implying that an amount of chance variation is superimposed on the main trend line of the response. It may clarify the point further if we contrast stochastics with the case of the fractal in which regular patterns arise from pure chaos. They may have seen response curves like the one in I1, but readers may not remember that these graphs show response shape and not a frequency distribution.

The consequence of the responses not being linear is a non-linear joint scatter as seen in Fig. I2

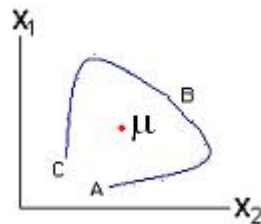


Fig. I2. Trend line in the presence of joint response of two species without stochastic oscillations for which the response graphs are shown in Fig. I1. Symbols X_1 , X_2 , A, B, C have identical meaning as in Fig. I1. Symbol μ denotes the centroid.

Figs. I1, I2 picture the case of only two populations (X_1 , X_2) exhibiting pure Groenewoud-Whittaker type response along a dominant environmental gradient. The gradient is long, and competition is absent. The populations' joint response depicted in Fig. I2 has a horseshoe shape. Multi-modal responses are not

uncommon in the presence of competition, in which case the joint scatter could be quite different from case to case under natural conditions.

The consequences of having a horseshoe-like joint scatter, and possibly even more complex non-linearity in the presence of many species for which the joint scatter would have to be drawn as some spiralling manifold in higher dimensions, can be devastating on the reliability of traditional statistical techniques of inference. The reason is, in part, their reliance on the moments-product moment and the p -variate Normal distribution with density function $f(X_1, X_2, \dots, X_p) = Be^{0.5(\mathbf{X}-\boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\mathbf{X}-\boldsymbol{\mu})}$. Since the product moment is useless to describe the non-linear relationship of the variables, the normal density function itself is rendered meaningless (see Orłóci 1993).

Choices have to be made and the question is whether to choose the centroid (centre of gravity, arithmetic mean vector as in Fig. I1) or the type (such as in preferential sampling) as the basis of prediction. Traditional statistics is interested in the centroid and it bases all inferences about populations on the 'centroid' in the sense of $E(\bar{X}) = \boldsymbol{\mu}$ with sample value on the left and its expectation on the right. When the sample consists of typical units, the behaviour of the 'type' is of interest. Will the behaviour of the centroid still have utility? The centroid could replace the 'type' if the responses were tested and found to be linear (Orłóci 1993). Should an actual case fall outside this specific type of response, such as in the aforementioned horseshoe configuration, the centroid will likely reside within a void where no natural state may ever materialise. In such a case, the centroid is the least likely state and it cannot be used to represent the 'type'.

It may be useful, with intentions to further clarify the notion of non-linearity and the problems with it, if we cast non-linearity in a broader context. Nayhef (2000) makes his point regarding non-linear systems an interesting way: the (geometric) principle of superposition does not hold and the source of non-linearity is the governing equation. To us engineers, it is quite natural to put the definition of non-linearity in these terms for reasons related to the non-linear relationship of stress to strain upon which the stability of a structural element, like a concrete beam, will depend. Such a beam normally contains steel, gravel, and cement in just the right proportions with a specific shape that experiments suggested. One of the peculiarities in the engineer's scenario is that the system is designed and the design does not violate the criterion equation(s) that are *a priori* chosen. Although stress and strain involves all the constituent materials of the beam, the total stress is not reproducible by superposing the individual stress and strain relations individually in the component parts: steel, gravel and cement. In other words, the non-linear system cannot

be reproduced like a montage of its constituent parts. Dawkins' (1986) makes this idea vivid when he makes a distinction between a heap of aeroplane parts and the aeroplane that flies.

In the ecologist's study scenario, the nonlinear system is a given, designed by nature. The governing equation is unknown and finding it, or using a less mathematical term, discovering the governing principles of the system, is the ultimate objective. It is exactly for the untenable principle of superposition that reductionism has utterly failed in ecology in its attempts to reveal governing principles of community dynamics. This is of course not to imply that the effort was wasted in other respects, such as the discovery of new population and individual behaviour patterns when populations are under stress.

Nash and Walker-Smith (1987) give simple examples, which reflect their view of non-linearity and its statistical analysis, non-linear parameter estimation:

- relationship of plant yield and the density of planting;
- change over time of the size of plants, animals or economic activities, *i.e.*, growth curves;
- rate of production or disappearance of a chemical or radioisotope in a chemical or nuclear reaction.

These are examples of non-linear responses (effects, stresses) to factor influences (causes, strains), be it the planting density or something analogous in manner, but not necessarily in functional form, to the case pictured in Figs. I1, I2 in the present essay. Nash and Walker-Smith (1987) see the statistical problem amounting to parameter estimation, but the scenario they outline is based on the assumption of the functional form of response(s) and the form of distribution of the error (noise, random oscillations) component about the response function. Their model is in fact regression analysis, the most used and abused statistical technique ever devised, for which the basic mathematical tools include Newtonian minimisation and Gaussian least squares.

It is relevant that regression tools have been used successfully in the modelling of complex ecological problems, such as the modelling of dynamic systems (Goodall 1967 and Wildi 1978, 1998, 2000). It is also true, of course, that in some ways successful models incorporate governing or state equations that the modellers construct to conform to educated guesses of the non-linear process, which they then improve in the course of successive approximation. So important and revealing is the fact that these equations are not the shorthand expressions of all encompassing truths, like Newton's or Einstein's, which as Mees (2001) points out are an impossibility to achieve in the circumstances of a 'real systems' on this planet that are noisy, non-stationery and having high-dimensional dynamics, such as in Goodall's paddock, Wildi's (1978) peat bog, our heathland and rain-forest examples. In characterising this kind of modelling, Mees is adamant that Kepler's approach is the more relevant model, because it works from masses of data and thus in that

manner reconstructs the governing equation. The ecologist approach, clearly, must avoid the mistake of committing itself to and staying with a preconceived idea about dynamics, such as for example the assumption that basic behaviour of change is a derivative of the exponential function, which was proven not to be the case in all but the most trivial type of communal existence.

Prediction: the method of analogues

The reader is referred to examples for the use of 'analogues' by Box (1981), Kűchler (1974) and Aszalós and Horváth (1998).¹ The method assumes that survival of specific plant forms in a site is not a straight chance event, and prediction is inherently an exercise in making choices among probable outcomes. But the outcome chosen for being most probable also has a certain probability of not being the one that actually materialises.² The *modus operandi* of predicting, say, community diversity as an outcome may go like this:

- a. Determine probabilistic linkage $p_{t|e}$ of each community type t to a specified type of the environment e .
- b. Community type t that has highest probability $p_{t|e}$ with regard to environmental type e , that has the highest probability p_e of occurring at a future point of time on a specific site, or at a different site in the same time, is the predictor community. The 'diversity' in the predictor community is the prediction. The probability of the prediction being right is $p_{t|e}p_e$.

Testing generality

Conclusions that have high value to science at large have validity broader than the data set itself. Such a conclusion of course has generality. How to test for generality? Two methods:

- a. Statistical testing. We have to emphasise the importance of the sampling design by the use of which data are collected, and the analytical methods by which the conclusions are drawn from the data. We discuss aspects of sampling in APPENDIX J. As far as statistical testing is concerned, basic conjectures called 'hypotheses' are formulated and then the question asked: what is the probability that a case specified by the hypothesis can come about by chance alone. If the probability is large, the hypothesis is validated, otherwise it is falsified. The technique requires a test criterion by which we can measure how far an observation has fallen from the hypothesised state, a probability distributions that describe the behaviour

¹ Other predictive models are used as well in ecology. Particularly good example is the Markov chain for which the reader finds examples and sources listed in Appendix B..

² Chance interplaying with the deterministic laws of adaptation and ecological plasticity may be in the manner of randomness infesting determinism. This is a case of stochastics and it has to be treated as such (Orlóci 2001a). It is not inconceivable that determinism could arise from complete randomness, the case of the fractal (Mandelbrot 1977), albeit not a likely mechanism on which to base the practice of prediction.

of the test criterion under the assumption that the hypothesis tested is true, and a decision rule to be invoked when rejecting (or accepting) a hypothesis. Given $\hat{\theta}$, an observed state, and θ_o , a state we would expect under a hypothesis H_o whose validity we wish to test, the test criterion may take the form of $\Delta = \hat{\theta} - \theta_o$. If we find that Δ is unusually large, so much so that it could not be expected to occur as a low probability event under the null hypothesis, then we would consider Δ significant and the H_o rejected. We would consider Δ having reached rejection limit when the probability of a difference at least as large as the observed Δ occurring by chance and chance alone when the null hypothesis is true is smaller than a set value. This probability has to be found. It may come from considerations of basic principles, a method standard in statistics of the classical type, or more likely, it may come from Monte Carlo experiments on which we have commented earlier in the text.

2. Testing by analogues. What we just described is a rather formal way of approaching the validation of a conjecture. More likely, one will follow a less formal approach where significance is linked to the consistency of similar results reoccurring in repeated surveys or experiments. In this, the idea of recursive sampling and analysis are key ideas (APPENDIX J). Significance is declared if a stable $\Delta = \hat{\theta} - \theta_o$ is found to be unusually large, consistent with expert experience.

APPENDIX J. Sampling

J1. Contrasting views of the sampling environment

Science is done by way of surveys and experiments, contemplative or applied. Applications tend to rely on samples, rarely on complete enumeration. The conceptualisation of the sampling environment is far from being uniform (Orlóci 1993). In fact, conceptualisations divide along a sharp dichotomy with traditional statistics on one side and ecological statistics on the other:

- a. The traditional statistics assumes that the sampling environment is homogeneous or can be made that way. It also holds that the sampling units are crisply defined and have unhindered access when chosen, the linear form of joint response is universal. Inference can become a matter of tracing the behaviour of moments and product moments, and probability scaling is possible based on theoretical sampling distributions derivable (axiomatically) from first principles.
- b. From an ecologist's point of view these assumptions should be tantamount to self-deception. It is a fact of the ecological environment that heterogeneity is the norm and not the exception; the population units

are fuzzy, access to sampling units is far from being unhindered, joint responses are never linear, and probability distributions cannot be derived from first principles.

J2. Unit selection: random or preferential?

We take samples since we cannot enumerate all population units. Which method of sampling should we choose? Unfortunately what could be an easy way out, namely random sampling, is not a viable option (Pielou 1977, Edgington 1987), except in the most trivial cases. This leaves us with the only practical method left: preferential sampling (Orlóci 1993). We replace letting chance govern the choice with selecting the type. The result is a sample in which the units are typical specimens, typical sola, typical community stands, or typical things of some other kind. After having said these, let the statistician in us be comforted by further suggesting that much of what is known about the biota on all scales from life zone to molecule, and about the environment on all scales from Ice Age climates to pedogenic processes, is the consequence of smart people looking closely at materials that they considered suitable for analysis, rather than having bothered working with materials that blind chance hands to them.

J3. Stopping rule in sampling: variance based or structural stability based?

Two criteria are particularly relevant when we formulate the stopping rule in sampling (Orlóci and Pillar 1989). The sampling variance written in the least cumbersome manner as in Sampford (1962), $SV = \frac{V}{n} \left(1 - \frac{n}{N}\right)$ incorporates all the essentials we need, to see what is happening when the stopping rule is SV based. In this n is sample size, N the population size, V the population variance. Clearly, the larger is n , the smaller will be SV . The SV graph is a rapidly descending curve which levels off early if N is reasonably large (Fig. J1). The sampling process stops when an *a priori* chosen sampling intensity is reached.

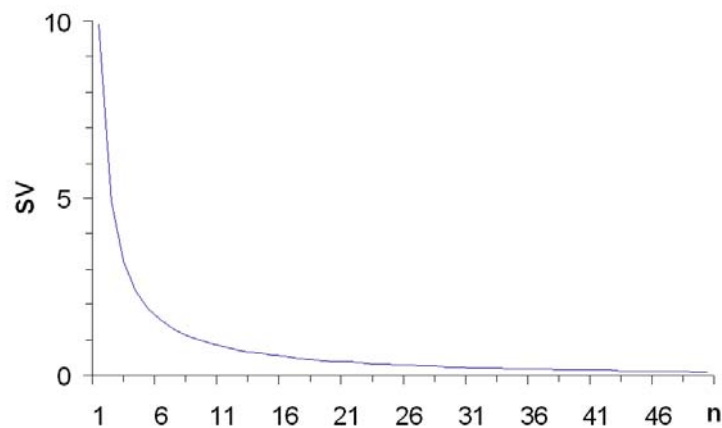


Fig. J1. The graph SV as a function of sample size n .

The fatal problem with the use of SV is that there is nothing inherently ecological about SV and even if there were its use would be hindered by V or N not being known in all but the most trivial cases in ecology. Beyond these, ecological studies are interested in structures and structural connections for which reason they will do better when the stopping rule is formulated about the stability of structures and structural connections in the sample (Orlóci and Pillar 1989). What is involved? Examination of Fig. J2 should help clarify this point. Suppose that two structural descriptions of the same n sampling units are given in the manner of two distance configurations. The first is $\mathbf{D}_{n \times n}$ based on vegetation data and the second $\mathbf{A}_{n \times n}$ based on environmental data. The signal for optimal sample size n having been reached is the levelling off of the stress function $\sigma(\mathbf{D}; \mathbf{A})$. It is noted that even though $\sigma(\mathbf{D}; \mathbf{A})$ is a measure of stress, its actual magnitude will reflect on the strength of linkage between \mathbf{D} and \mathbf{A} and will not affect the determination of the optimal sample size.

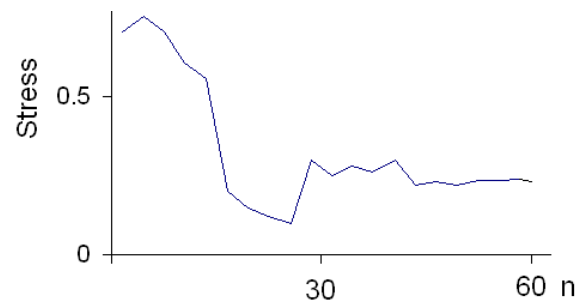


Fig. J2. Stress graph for distance configurations \mathbf{D} and \mathbf{A} , corresponding to vegetation structure and environmental structure in the sample according to Orlóci and Pillar (1989). \mathbf{D} and \mathbf{A} change in the sample by addition of new units. They attained reasonably high stability in the example by the time the sample size reached 30. Note, low stress in terms of $\sigma(\mathbf{D}; \mathbf{A})$ implies high concordance of \mathbf{D} and \mathbf{A} .

REFERENCE LIST

(for main text and Web Only Appendices)

- Anand M (1994). Pattern, process and mechanism. The fundamentals of scientific inquiry applied to vegetation science. *Coenoses* 9: 81-92.
- Anand M (2000). The fundamentals of vegetation change: complexity rules. *Acta Biotheoretica* 48:1-14.

- Anand M, Orlóci L (1997). Chaotic dynamics in a multispecies community. *Environmental and Ecological Statistics* 4: 337-344.
- Anand M, Orlóci L (1996). Complexity in plant communities: the notion and quantification. *J. theor. Biol.* 179: 179-186.
- Anand M, Orlóci L (2000). On partitioning of an ecological complexity function. *Ecological Modelling* 132: 51-62.
- Anderson G (1896). *Swenska vaextvaerldens*. Stockholm.
- Anderson P M (1988). Late Quaternary pollen records from the Kobuk and Noatak River drainages, northwestern Alaska. *Quaternary Research* 29: 263-276.
- Anderson R S (1993). A 35,000 year vegetation and climate history from Potato Lake, Mogollon Rim, Arizona. *Quaternary Research* 40:351-359.
- Aszalós R, Horváth F (1998). Prediction of vegetation pattern on the regional scale. (In Magyar.) In: Fekete, G. ed. *The Front Lines of Community Ecology*, pp. 161-170. Scientia, Budapest.
- Bartha S, Czárán T, Scheuring I (1998). Spatio-Temporal scales of non-equilibrium community dynamics in abstract coenostate spaces. *Abstracta Botanica* 22: 49-66.
- Behling H, De Patta Pillar V, Orlóci L, Bauermann S G (2004). Late Quaternary Araucaria forest, grassland (Campos), fire and climate dynamics, studied by high-resolution pollen, charcoal and multivariate analysis of the Cambará do Sul core in southern Brazil. *Paleogeography, Paleoclimatology, Paleoecology* 203: 277-297.
- Birks H J B, Gordon A D (1985). *Numerical Methods in Quaternary Pollen Analysis*. Academic Press, New York.
- Blunier T, Brook E J (2001). Timing of millennial-scale climate change in Antarctica and Greenland during the Last Glacial Period. *Science* 291: 109-112.
- Bonnefille R, Riollet G, Buchet G, Icole M, Lafont R, Arnold M, Jolly D (1995). Glacial/Interglacial record from intertropical Africa, high resolution pollen and carbon data at Rusaka, Burundi. *Quaternary Science Reviews* 14: 917-936.
- Box E. O (1981). *Macroclimate and Plant Forms: An Introduction to Predictive Modelling in Plantgeography*. W. Junk bv, The Hague.
- Braun E L (1950). *Deciduous Forests of Eastern North America*. Blakston, Toronto.
- Braun-Blanquet J (1927). *Pflanzensoziologie*. Montpellier, France.
- Braun-Blanquet J (1932). *Plant Sociology*. (Translated by G.D. Fuller and H.S. Conard). McGraw-Hill, New York.
- Cajander A K (1909). Über Waldtypen. *Acta Forestalia Fennica* 1, Helsinki.
- Çambel A B (1993). *Applied Chaos Theory: a Paradigm for Complexity*. Academic Press, New York.
- Cain S A (1944). *Foundations of Plant Geography*. Harper, New York.
- Clements F E (1916). *Plant Succession: an Analysis of the Development of Vegetation*. Publ. No. 242, Carnegie Institution, Washington.
- Colinvaux P A , de Oliveira P E, Moreno J E, Miller M C, Bush M B (1996). A long pollen record from lowland Amazonia: forest and cooling in glacial times. *Science* 274: 85-88.
- Conard H S (1951). *The Background of Plant Ecology*. The Iowa State University Press, Ames. --- Note: This is a translation of Kerner von Merilaun's "Das Pflanzenleben der Donauländer" first published in 1863. See also main entry for Kerner in this list of references. New printing was marketed by Arno Press, New York 1977.
- Connell J H, Slatyer R O (1977). Mechanisms of succession in natural communities and their role in community stability and organization. *Am. Nat.* 111: 1119-1144.
- Cowles H C (1899). The physiographic ecology of Chicago and vicinity: a study of the origin, development, and classification of plant societies. *Bot. Gaz.* 31: 73-108, 145-82.
- Cox D R, Lewis P A W (1968). *The Statistical Analysis of Series of Events*. Methuen, London.

- Cwynar L C (1982). A Late-Quaternary vegetation history from Hanging Lake, northern Yukon. *Ecological Monographs* 52: 1-24.
- Czárán T (1998). *Spatiotemporal Models of Population and Community Dynamics*. Vol. 21. Population and Community Biology Series. Chapman and Hall, London.
- Czárán T, Bartha S (1989). The effect of spatial pattern on community dynamics: a comparison of simulated and field data. *Vegetatio* 83: 229-239.
- Dansereau P (1951). *Biogeography, an ecological perspective*. Ronald Press, New York.
- Dawkins R (1986). *The Blind Watchmaker*. Penguin Books, Suffolk.
- Delcourt P A, Delcourt H R (1987). *Long-term Forest Dynamics of the Temperate Zone*, Springer-Verlag, New York.
- Diggle P J (1981). *Statistical Analysis of Point Patterns*. Academic Press, London.
- Dokučajev V V (1899). *Kučeniju o zonah prirody. Gorizontannyje I vertikalnyje počvennyje zony*. St. Petersburg.
- Edgington E S (1987). *Randomization Tests*. 2nd ed. Marcel Dekker, New York.
- Edwards A W F, Cavalli-Sforza L L (1965). A method for cluster analysis. *Biometrics* 21: 362-375.
- Egler F E (1954). Vegetation science concepts. I. Initial floristic composition - a factor in old-field vegetation development. *Vegetatio* 4: 412-417.
- Fekete G (1985). Terrestrial vegetation succession: theories, models, reality. In: G. Fekete ed. *Basic questions of Coenological Succession*, 31-63. In Magyar.) Akadémiai Kiadó, Budapest.
- Fekete G, Virágh K, Aszalós R, Orlóci L (1998). Landscape and ecological differentiation of *Brachypodium pinnatum* grasslands in Hungary. *Coenoses* 13: 39-53.
- Feller W (1957). *An introduction to probability theory and its applications*. Vol. I, Wiley and Sons, London.
- Feoli E, Orlóci L (1985). Species dispersion profiles of anthropogenic grasslands in the Italian Pre-Alps. *Vegetatio* 60: 113-118.
- Fewster H P L, Orlóci L (1983). On choosing a resemblance measure for non-linear predictive ordination. *Vegetatio* 54: 27-35.
- Fosberg F R (1965). The entropy concept in ecology. In: *Symposium on Ecological Research in Humid Tropics Vegetation*, pp. 157-163. UNESCO and Government of Sarawak, Kuching, Sarawak.
- Gleason H A (1926). The individualistic concept of the plant association. *Bull. Torrey Bot. Club* 53: 7-26.
- Gleason H A (1917). The structure and development of the plant association. *Bull. Torrey Bot. Club* 43.
- Gleick J (1987). *Chaos. Making a New Science*. Penguin Books, New York.
- Goodall D W (1954). Objective methods for classification of vegetation. III. An essay in the use of factor analysis. *Aust. J. Bot.* 2: 304-324.
- Goodall, D.W. 1967. Computer simulation of changes in vegetation subject to grazing. *J. Ind. Bot. Soc.* 46:56-61.
- Goodall D W (1972). Building and testing ecosystem models. In: Jeffers, N.R. (ed.), *Mathematical Models in Ecology*, pp. 173-214. Blackwell, Oxford.
- Greig-Smith P (1952). The use of random and contiguous quadrats in the study of the structure of plant communities. *Annals of Botany* 16: 293-316.
- Greig-Smith P (1957). *Quantitative Plant Ecology*. Butterworths, London.
- Greig-Smith P (1983). *Quantitative Plant Ecology*. 3rd ed. Blackwell Scientific, London.
- Grime J P (1977). Evidence for the existence of three primary strategies in plant and its relevance to ecological and evolutionary theory. *Am. Nat.* 111: 107-1169-1194.

- Györfly Gy, Zólyomi B (1996). The Carpathian Basin and Etelköz as they appeared one thousand years ago. (In Magyar.) Magyar Tudomány Vol. XLI, 8: 899-918.
- Hammersley J M, Handscom D C (1967). Monte Carlo Methods. Methuen, London.
- Harper J L (1977). Population Biology of Plants. Academic Press, London.
- He X S, Orlóci L (1999). Anderson Pond revisited: the late Quaternary vegetation process. *Abstracta Botanica* 22: 81-93.
- Hill M O, Gauch H G (1980). Detrended correspondence analysis: an improved ordination technique. *Vegetatio* 42: 47-58.
- Hotelling H (1933). Analysis of a complex of statistical variables into principal components. *J. Ed. Psych.* 24: 417-441, 498-520.
- Hult R (1881). Forsök til analytiskbehandling af växformationerna. *Medd. Soc. Faun. Flor. Fenn.* 8. --- Note: 1885. Blekinges vegetation. Ett bidrag till växformationernas. *Medd. Soc. Faun. Flor. Fenn.* 8.
- Hulst R van (1992). From population dynamics to community dynamics: modelling succession as a species replacement process. In: Glenn-Lewin, D.C., Peet, R.K., Veblen T T eds. *Plant Succession: Theory and Prediction*, pp. 188-214. Chapman and Hall, London.
- Hulst R van (2000). Vegetation dynamics and plant constraints: separating generalities and specific. *Community Ecology* 1: 5-12.
- Humboldt A von, Bonpland A (1807). *Ideen zu einer Geographie der Pflanzen nebst einer Naturgemälder Tropenländer*. F.G. Cotta, Tubingen.
- Humboldt L von (1850). *Views of Nature*. H. G. Bohn, London.
- Izsák J, Juhász-Nagy P, Varga Z (1982). *Introduction to Biomathematics*. (In Magyar.) 2nd ed. Tankönyv Kiadó, Budapest.
- Jacobs B F (1985). A middle Wisconsin pollen record from Hay Lake, Arizona. *Quaternary Research* 24: 121-130.
- Jeffers J N R (1978). *An Introduction to Systems Analysis: with Ecological Applications*. Edward Arnold, London.
- Juhász-Nagy P., Podani, J. 1983. Information theory methods for the study of spatial processes and succession. *Vegetatio* 51: 129-140.
- Juhász-Nagy P (1985). Introduction to syndynamics. (In Magyar.) In: Fekete G, ed. *Basic questions of Coenological Succession*, pp. 13-30. Akadémiai Kiadó, Budapest.
- Kenkel N C, Walker D J (1993). Fractals and ecology. *Abst. Bot.* 17: 53-70.
- Kerner von Marilaun A (1863). *Das Pflanzenleben der Danauländer*. Innsbruck, Wagner.
- Kershaw A P (1994). Pleistocene vegetation of the humid tropics of northeastern Queensland, Australia. *Palaeogeography, Palaeoclimatology, and Palaeoecology* 109: 399-412.
- Kovács-Láng E, Fekete G, Molnár Zs (1998). Pattern, process, scale: long-term ecological researches in Kiskunság. (In Magyar.) In: G. Fekete ed. *Frontiers of Community Ecology*. Scientia, Budapest.
- Krajina V J (1963). Biogeoclimatic zones on the Hawaiian Islands. *Newsletter of the Hawaiian Botanical Society* 7: 93-98.
- Kullback S (1959). *Information theory and Statistics*. Wiley, New York.
- Kullback S, Kuperman M, Ku H H (1962). Test for contingency tables and Markov chains. *Technometrics* 4: 573-608.
- Küchler A W (1974). *Vegetation Mapping*. Ronald Press, New York.
- Kühler A W (1990). Natural vegetation. In: Espenchiede Jr, E B, Morrison J L, eds. *Rand McNally Good's World Atlas*, 18th ed, pp. 8-9. Rand McNally, New York.
- Lagerheim G (1902). Metoder för pollenubder sökning. *Bot. Notiser* 75-78.
- Legendre P, Legendre L (1983). *Numerical Ecology*. Elsevier, Amsterdam.

- Legendre P, Anderson M J (1999). Distance based redundancy analysis: testing multispecies responses in multifactorial ecological experiments. *Ecological Monographs* 69: 1-24.
- Levin S A, ed. 2001. *Encyclopaedia of Biodiversity*. Academic Press, San Diego.
- Libby W F (1955). *Radiocarbon Dating*. 2nd ed. University of Chicago Press, Chicago.
- Lippe E, De Smidt J T, Glen-Lewin D C (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.
- Lotka A J (1925). *Elements of Physical Biology*. Williams and Wilkins, Baltimore.
- Lowe J J, Walker M J C (1997). *Reconstructing Quaternary Environments*. 2nd ed. Longman, Hong Kong.
- Lozhkin A V, Anderson P M, Eisner W R, Ravako L G, Hopkins D M, Brubaker L B, Colinvaux P A, Miller M C (1993). Late Quaternary lacustrine pollen records from southwestern Beringia. *Quaternary Research* 39: 314-324.
- Magurran A E (1988). *Ecological Diversity and Its measurement*. Croom Helm, London.
- Manabe S, Bryan K, Spelman M J (1990). Transient response of a global ocean-atmosphere model to a doubling of atmospheric carbon dioxide. *J. Phys. Oceanogr.* 20: 722-749.
- Manabe S K, Bryan - Spelman M J (1990). Transient response of a global ocean-atmosphere model to a doubling of atmospheric carbon dioxide. *J. Phys. Oceanogr.* 20: 722-749.
- 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.
- Mandelbrot B B (1983). *The fractal geometry of Nature*. Freeman, San Francisco.
- Margalef R (1989). On diversity and connectivity as historical expressions of ecosystems. *Coenoses* 4: 121-126.
- Mason J (1990). *The greenhouse effect and global warming*. Information Office, British Coal, C.R.E. Stoke Orchard, Cheltenham, Gloucestershire, U.K. GL52 4RZ.
- May R M ed. (1981). *Theoretical Ecology*. Blackwell, Oxford.
- May R M (1976). Simple mathematical models with very complicated dynamics. *Nature* 261: 459-471.
- May R M (1987). Chaos and the dynamics of biological populations. *Proc. Royal Soc. London, Series A* 413: 27-44.
- May R M, Oster G F (1976). Bifurcations and dynamic complexity in simple ecological model. *The American Naturalist* 110: 573-599.
- McArdle, B. H., Anderson, M. J. 2001. Fitting multivariate models to community data: a comment on distance based community analysis. *Ecology* 82: 290-297.
- McIntosh R P (1967). The continuum concept of vegetation. *Bot. Rev.* 33: 130-187.
- McIntosh R P (1985). *The Background of Ecology: Concept and Theory*. Cambridge University Press, New York.
- Mees A I ed. (2001). *Nonlinear Dynamics and Statistics*. Birkhäuser, Boston.
- Metropolis N, Ulam S (1949). The Monte Carlo method. *Journal of the American Statistical Association* 44:335-341.
- Metropolis N (1987). The beginning of the Monte Carlo method. *Los Alamos Science Special Issue*, pp. 125-130. <http://library.lanl.gov/cgi-bin/getfile?00326866.pdf>.
- Milankovitch M M (1941). Canon of isolation and the Ice-Age problem. *Royal Serb Academy Special Publication* 133.
- Morozov G F (1912). *Ucenije o Lese*. Leningrad. Note other sources: -- (1928). *Die Lehre vom Walde*. Neudamm, J. Neumann. -- (1952). *Az Erdő Élettana*. Budapest, Mezőgazdasági Kiadó.

- Morrison D F (1976). *Multivariate Statistical Methods*. McGraw-Hill. London.
- Mueller-Dombois D (1992). A natural dieback theory, cohort senescence, as an alternative to the decline disease theory. In: P.D. Manion and D. Lachance, eds. *Forest Decline Concepts*, pp. 26-37. The Am. Phytopath. Soc., St. Paul, Min.
- Nash J C, Walker-Smith M (1987). *Nonlinear Parameter Estimation. An Integrated System in BASIC*. Marcel Dekker, New York.
- Nayfeh A H (2000). *Nonlinear Interactions. Analytical, Computational, and Experimental Methods*. Wiley-Interscience, New York.
- Orlóci L (1966). Geometric models in ecology I. The theory and application of some ordination methods. *J. Ecol.* 54: 193-215.
- Orlóci L (1967). Data centering: a review and evaluation with reference to component analysis. *Syst. Zool.* 16: 208-212.
- Orlóci L (1974). On information flow in ordination. *Vegetatio* 29: 11-16.
- Orlóci L (1978). *Multivariate Analysis in Vegetation Research*. W. Junk bv, The Hague.
- Orlóci L (1991a). On character-based community analysis: choice, arrangement, comparison. *Coenoses* 6: 103-107.
- Orlóci L (1991b). Entropy and Information. *Ecological Computations Series (ECS) Vol. 3*. SPB Academic Publishing bv, the Hague.
- Orlóci L (1991c). CONAPACK: A program for Canonical Analysis of Classification Tables. *Ecological Computations Series: Vol. 4*. SPB Academic Publishing, The Hague.
- Orlóci L (1993). The complexities and scenarios of ecosystem analysis. In: Patil, G.P., Rao, C.R. eds. *Multivariate Environmental Statistics*, pp. 423-432. Elsevier Scientific, New York.
- Orlóci L (1994). Global warming: the process and its anticipated phytoclimatic effects in temperate and cold zone. *Coenoses* 9: 69-74.
- Orlóci L (2000). From Order to Causes. A personal view, concerning the principles of syndynamics. Internet essay. <http://publish.uwo.ca/~lorloci/files.html>.
- Orlóci L (2001a). Pattern dynamics: an essay concerning principles, techniques, and applications. *Community Ecology* 2: 1-15.
- Orlóci L (2001b). Prospects and expectations: reflections on a science in change. *Community Ecology* 2: 187-196.
- Orlóci L (2006). Diversity partitions in 3-way sorting: functions, Venn diagram mappings, typical additive series, and examples. *Community Ecology* 7: 253-259.
- Orlóci L (2008). Vegetation displacement issues and transition statistics in climate warming cycle. *Community Ecology* 9:83-98.
- Orlóci L, Orlóci M (1985). Comparison of plant communities without the use of species: model and example. *Annali di Botanica (Roma)* 43: 275-285.
- Orlóci L, Orlóci M (1988). On recovery, Markov chains and canonical analysis. *Ecology* 69: 1260-1265.
- Orlóci L, Orlóci M (1990). Edge detection in vegetation: Jornada revisited. *Journal of Vegetation Science* 1: 311-324.
- Orlóci L, Orlóci M (1995). *Sampling and Data Analysis. Theory, problems, examples*. Mimeographed Lecture Notes, SCADA Associates, London.
- Orlóci L, De Patta Pillar V (1989). On sample size optimality in ecosystem survey. *Biometrie-Praximétrie* 29: 173-184.
- Orlóci L, Anand M, He X S (1993). Markov chain: a realistic model for temporal coenoser? *Biométrie-Praximétrie* 33: 7-26.
- Orlóci L, Pillar V D, Anand M, Behling H (2002a). Some interesting characteristics of the vegetation process. *Community Ecology* 3: 125-146.

- Orlóci, L., Pillar, V. D., Anand, M. 2006. Multiscale analysis of palynological records: new possibilities. *Community Ecology* 7:53-68.
- Orlóci, L., Anand, M., Pillar, V.D. 2002b. Biodiversity analysis: issues, concepts, techniques. *Community Ecology* 3: 217-236.
- Palmer M W (1988). Fractal geometry: a tool for describing spatial pattern of plant communities. *Vegetatio* 75: 91-102.
- Palmer M W (1992). The coexistence of species in fractal landscapes. *Am. Nat.* 139: 375-397.
- Patil G P, Taillie C (1979). An overview of diversity. In: Grassle J F, Patil G P, Smith W, Taillie C eds. *Ecological Diversity in Theory and Practice*, pp. 3-27. International Co-operative Publishing House, Fairland, Maryland.
- Peet R K (1974). The measurement of species diversity. *Annual Review of Ecology and Systematics* 5: 285-307.
- Petit J R., Jouzel D, Raynaud D, Barkov N I, Barnola J M, Basile I, Bender M, Chappellaz J, Davis J, Delaygue G, Delmotte M, Kotlyakov V M, Legrand M, Lipenkov V, Lorius C, Pepin L, Ritz C, Saltzmann E, Stievenard M (1999). Climate and atmospheric history of the past 420,000 years from the Vostok Ice Core, Antarctica. *Nature* 300: 429-436.
- Petit J R, Raynaud D, Lorius C, Jouzel J, Delaygue D, Barkov N I, Kotlyakov V M (2000). Historic isotopic temperature records from the Vostok Ice Core. In: *Trends: A compendium of Data on Global Change. Carbon Dioxide Information Analysis Center, Oak Ridge National Laboratory, U.S. Department of Energy, Oak Ridge, Tenn., U.S.A.*
- Petit J R., Jouzel D, Raynaud D, Barkov N I, Barnola J M, Basile I, Bender M, Chappellaz J, Davis J, Delaygue G, Delmotte M, Kotlyakov V M, Legrand M, Lipenkov V, Lorius C, Pepin L, Ritz C, Saltzmann E, Stievenard M (2001). Vostok Ice Core Data for 420,000 years, IGBP PAGES/World Data Center for Paleoclimatology Data Contribution Series 2001-076. NOAA/NGDC Paleoclimatology Program, Boulder CO, USA
- Phillips J (1935). Succession, development, the climax and the complex organism: an analysis of concepts. Part III. *J. Ecol.* 23: 488-508.
- Pielou E C (1977). *An Introduction to Mathematical Ecology*. Wiley-Interscience, New York.
- Pielou E C (1975). *Ecological Diversity*. Wiley, New York.
- Pillar V D, Orlóci L (1993a). *Character-based Vegetation Analysis: the Theory and an Application Program*. Ecological Computations Series (ECS): Vol. 5. SPB Academic Publishing bv, The Hague, The Netherlands.
- Pillar V D, Orlóci L (1993b). Taxonomy and perception in vegetation analysis. *Coenoses* 8: 53-66.
- Pillar V D, Orlóci L (1996). On randomisation testing in vegetation science: multifactor comparison of relevé groups. *Journal of Vegetation Science* 7: 587-592.
- Pillar V D, Sosinski Jr E E (2003). An improved method for searching plant functional types by numerical analysis. *Journal of Vegetation Science* 14: 323-332.
- Pillar V D, Orlóci L (2004). *Character-Based Community Analysis: The Theory and an Application Program*. Electronic Edition, available at <http://ecoqua.ecologia.ufrgs.br>.
- Pillar V D (1999). On the identification of optimal plant functional types. *Journal of Vegetation Science* 10: 631-640.
- Podani J (1994). *Multivariate Analysis in Ecology and Systematics*. SPB Publishing, The Hague.
- Poore M E D (1962). The method of successive approximation in descriptive ecology. In: *Advances in Ecological Research*. Vol. 1, pp. 35-68. Academic Press, New York.
- Post L von (1916). Forest tree pollen in south Swedish peat bog deposits. (In Swedish.) *Pollen et Spores* 9: 375-401.
- Post L. Von (1946). The prospect for pollen analysis in the study of the earth climatic history. *New Phytol.* 45: 193-217.
- Prigogine I, Stengers I (1984). *Order out of chaos*. Bantam Books, New York.
- Rajski C (1961). Entropy and metric spaces. In: Chery C, ed, *Information Theory*, pp. 41-45. Butterworths, London.
- Rao C R (1982). Diversity and dissimilarity coefficients: a unified approach. *Theor. Popul. Biol.* 21: 24-43.
- Reyner J N (1971). *An Introduction to Spectral Analysis*. Pion Limited, London.

- Rényi A (1961). On measures of entropy and information. In: Neyman J, ed. Proceedings of the 4th Berkeley Symposium on Mathematical Statistics and Probability, pp. 547-561. University of California Press, Berkeley.
- Ripley B D (1981). Spatial Statistics. Wiley & Sons, New York.
- Sampford M R (1962). An Introduction to Sampling Theory. Oliver & Boyd, Edinburgh.
- Scheuring I (1993). Multifractality: a new concept in vegetation science. *Abstracta Botanica* 17: 71-77.
- Schönemann P H, 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.
- Schweingruber F H (1996). Tree Rings and Environment Dendroecology. Swiss Federal Institute for Forests Snow and Landscape Research, Birmensdorf and Paul Haupt, Bern.
- Shchigolev B M (1960). Mathematical analysis of observations. Elsevier, New York.
- Shugart H H ed. 1977. Time Series and Ecological Processes. SIAM, Philadelphia.
- Shugart H H Jr. 1984. A Theory of Forest Dynamics: the Ecological Implications of Forest Succession Models. Springer-Verlag, New York.
- Simpson E H (1949). Measurement of diversity. *Nature* 163: 688.
- Singh G, Geissler E A (1985). Late Cainozoic history of vegetation, fire, lake levels and climate at Lake George, New South Wales, Australia. *Philosophical Transactions of the Royal Society of London Series B* 311:379-447.
- Sukachev V N (1913). Introduction to the Study of Plant Communities. (In Russian.) *Bibl. Natur. St. Petersburg*.
- Sukopp H (1987). On the history of plant geography and plant ecology in Berlin. *Englera* 7: 85-103.
- Tansley A G (1935). The use and abuse of vegetational concepts and terms. *Ecology* 16: 284-307
- Tóth-Mérész B (1997). Diversity Orderings. (In Magyar.) *Scientia Kiadó, Budapest*.
- Tóth-Mérész B (1998a). On the characterisation of scale-dependent diversity. *Abstracta Botanica* 22: 149-1956.
- Tóth-Mérész B (1998b). Quantitative ecological methods for examination of scale dependences. In: Fekete G, ed. *The Frontiers of Community Ecology*. *Scientia Kiadó, Budapest*, pp. 145-160. (In Magyar).
- Trewartha G T (1990). Climatic regions. In: Espenshade E B Jr., Morrison J L eds. *Rand McNally Good's World Atlas*, 18th ed. pp. 8-9. Rand McNally, New York,
- Trewartha G T (2001). Global Mechanism of UNCCD, Via del Serafico 107, 00142 Rome, Italy. Web address: www.gm-uncd.org/English/Field/aridity.htm.
- Usher M B (1981). Modelling ecological succession with particular reference to Markovian models. *Vegetatio* 46: 11-18.
- Usher M B (1992). Statistical models of succession. In: Glenn-Lewin D C, Peet R K, Veblen T T eds. *Plant Succession: Theory and Prediction*, pp. 215-248. Chapman and Hall, London.
- Volterra V (1926). Variation and fluctuation of the number of individual species living together. *J. Cons. Perm. Int. Ent. Mer.* 3: 3-51. Reprinted in: Chapman, R.N. 193. *Animal Ecology*. McGraw-Hill, NY.
- Walker D J, Kenkel N C (1999). Fractal analysis of spatio-temporal dynamics in boreal forest landscapes. *Abstracta Botanica* 22: 13-28.
- Walter H, Harnickell E, Mueller-Dombois D (1975). *Climate Diagram Maps*. Springer-Verlag, New York.
- Watt A S (1947). Pattern and process in the plant community. *J. Ecol.* 35: 1-22.
- Watts W A, Bradbury J P (1982). Paleocological studies at Lake Patzcuaro on the west-central Mexican Plateau and at Chalco in the Basin of Mexico. *Quaternary Research* 17: 56-70.
- Watts W A, Hansen B C S, Grimm E C (1992). Camel Lake: A 40,000-yr record of vegetational and forest history from northwest Florida. *Ecology* 73: 1056-1066.

- Whittaker R H (1953). A consideration of climax theory: the climax as a population pattern. *Ecol. Monogr.* 23: 41-78.
- Whittaker R H (1962). Classification of natural communities. *Bot. Rev.* 28: 1-239.
- Whittaker R H (1967). Gradient analysis of vegetation. *Bot. Rev.* 42: 207-264.
- Wildi O (1978). Simulation of development in peat bogs. *Vegetatio* 37: 1-17.
- Wildi O (1998). Simulating vegetation at a local scale. *Abstracta Botanica* 22: 3-11.
- Wildi O, Orlóci L (1991). Flexible gradient analysis: a note on ideas and an application. In: Feoli, E., Orlóci, L. (Eds), *Computer Assisted Vegetation Analysis*, pp. 249-254. Kluwer, Dordrecht.
- Wildi O, Orlóci L (2007). Essay on the study of the vegetation process. In: Kienast F, Wildi O, Ghosh S eds, *A Challenging World*, pp. 195-207. Springer, The Netherlands.
- Wildi O, Schütz M (2000). Reconstruction of a long-term recovery process from pasture to forest. *Community Ecology* 1: 25-32.
- Wilkins G R , Delcourt P A, Delcourt H R, Harrison F W, Turner M R (1991). Paleocology of central Kentucky since the last glacial maximum. *Quaternary Research* 36:224-239.