DETECTION OF ATYPICAL YEARS IN TREE-RING SERIES BY CONSTRUCTION OF A TEMPORAL WALK IN THE PRINCIPAL COMPONENTS PLANES

PIERRE DUTILLEUL

Unité de Biométrie et Analyse des Données
Faculté des Sciences Agronomiques
Université Catholique de Louvain
Louvain-la-Neuve, Belgium

and

CLAUDINE TILL

Rue des Blancs Chevaux 1/404
B-1348 Louvain-la-Neuve, Belgium

ABSTRACT

The purpose of this paper is to introduce a method for identifying years of anomalous radial growth, which are called atypical years and are characterized by particularly narrow rings for certain trees or sites, and wide for the others. With this aim, we use principal components analysis (PCA) of tree-ring series, where years are the variables and trees or sites are the individuals, even though in classical dendrochronological applications of PCA, the trees or sites are considered as variables and the years as cases. The relevant method is explained and results are given for five cedar forests (*Cedrus atlantica* (Endl.) Carrière) in Morocco.

In der vorliegenden Arbeit wird ein Verfahren vorgestellt, das es ermöglicht, Jahre mit anomalem Dickenwachstum zu identifizieren. Derartige Jahre werden als atypisch bezeichnet und fallen durch besonders enge Jahrringe für bestimmte Bäume oder Standorte und besonders breite Jahrringe für andere Bäume oder Standorte auf. Zur Indentifikation wird eine Hauptkomponentenzerlegung der Jahrringfolgen eingesetzt. Hierbei gelten die Jahre als Variable, und die Bäume oder Standorte sind die Individuen, obwohl bei der klassischen Anwendung der Hauptkomponentenzerlegung in der Dendrochronologie die Bäume oder Standorte als Variable und die Jahre als Beobachtungsfälle behandelt werden. Die angewandte Methode wird erklärt, und Ergebnisse für fünf Cedern-Wälder (*Cedrus atlantica* (Endl.) Carrière) in Marokko werden dargestellt.

L'objectif de cet article est d'introduire une méthodologie permettant d'identifier les années de croissance radiale anormale, qualifiées d'atypiques et caractérisées par des épaisseurs de cernes particulièrement petites pour certains arbres ou sites, et larges pour les autres. Nous appliquons dans ce but l'analyse en composantes principales (ACP) sur des chronologies d'épaisseurs de cernes, où les années sont les variables et les arbres ou sites sont les individus, alors que dans les applications dendrochronologiques classiques de l'ACP, les arbres ou sites sont considérés comme variables et les années comme observations. La méthodologie est expliquée et des résultats sont donnés pour cinq cédraies (*Cedrus atlantica* (Endl.) Carrière) au Maroc.

INTRODUCTION

Principal components analysis (PCA) is a classical method for data analysis (e.g., see Lebart and Fénelon 1975; Lebart *et al.* 1977, 1979). In dendrochronology, PCA has been applied directly to tree-ring series (Jacoby and Cook 1981; Peters *et al.* 1981) but has more generally been used in order to summarize meteorological data before computing response functions (Fritts 1976; Guiot 1981; Guiot *et al.* 1982).

In classical PCA of tree-ring series, trees or sites are considered as variables and years, as observations. Here, we analyze the transpose of such tree-ring series data arrays, *i.e.* years are taken as variables. The purpose is to detect years of anomalous radial growth, which are called *atypical years*. We define a year as atypical when tree rings are particularly narrow for certain trees or sites, and wide for the others. According to this definition, a year with generally narrow rings or a year with generally wide rings will not be considered atypical. In a discriminant analysis of 44 Moroccan cedar forests (Dutilleul and Till 1988), these atypical years contributed heavily to the construction of the two first canonical axes, which in turn clearly reproduced the geographical location of the forests and in particular, the proximity of the Rif and Western Middle Atlas sites.

TREE-RING DATA

Forty-four sites have been sampled across the whole natural area of the cedar forest (*Cedrus atlantica* (Endl.) Carrière) in Morocco. Five sites have been selected for the period 1860-1976 to perform the present analysis. These sites have been chosen according to their geographic locations and their tree-ring characteristics. Only five sites have been retained in order to allow easier interpretation of the PCA results. These sites are marked on the map (Figure 1). Sites 22 and 40 are from the Rif mountains, sites 34 and 47 from the Middle Atlas and site 5 from the Eastern High Atlas. Site 22 has been classified as complacent with mean sensitivity s=0.08 and first-order autocorrelation coefficient r=0.71; site 5 (s=0.27, r=0.50), site 40 (s=0.21, r=0.36) and site 47 (s=0.18, r=0.21) as intermediate; site 34 (s=0.30, r=0.09) as sensitive (Till 1987b).

According to the woody material available, 10 to 20 trees have been sampled in each site; 2 to 4 cores have been extracted from each tree. After measuring ring widths to the nearest hundredth of a millimeter and crossdating the ring-width series for all sites, trees were classified as young-adult stage or old-adult stage (Till 1985, 1987a, 1987b). Individual chronologies and master chronologies have been computed. An individual chronology is a mean chronology summarizing synchronized cores of an individual tree. A master chronology is an average of the chronologies relative to synchronized cores from the same site. For the present paper, only five or six old-adult stage trees (10 to 12 synchronized cores) have been randomly retained for each site, and the mean chronologies (individual and master chronologies) concern these trees only, extending over the period 1860-1976.

METHOD

Data Arrays and Data Transformations

Tree-ring width series relative to the sites under study may be arranged in data arrays

$$X = (X_{it})$$

where i denotes the site or the individual (i=1, ..., I) and t denotes the year (t=1860, ..., 1976); X_{it} is the ring-width of the ith individual or site for year t.

PCA is performed on the raw data arrays, and two transformations are studied for their effects on PCA results. These transformations are the following:

1. centering of the tree-ring series (X rows):

$$X^c = (X^c_{it})$$

where:

$X^c_{it} = X_{it} - X_{.t}$

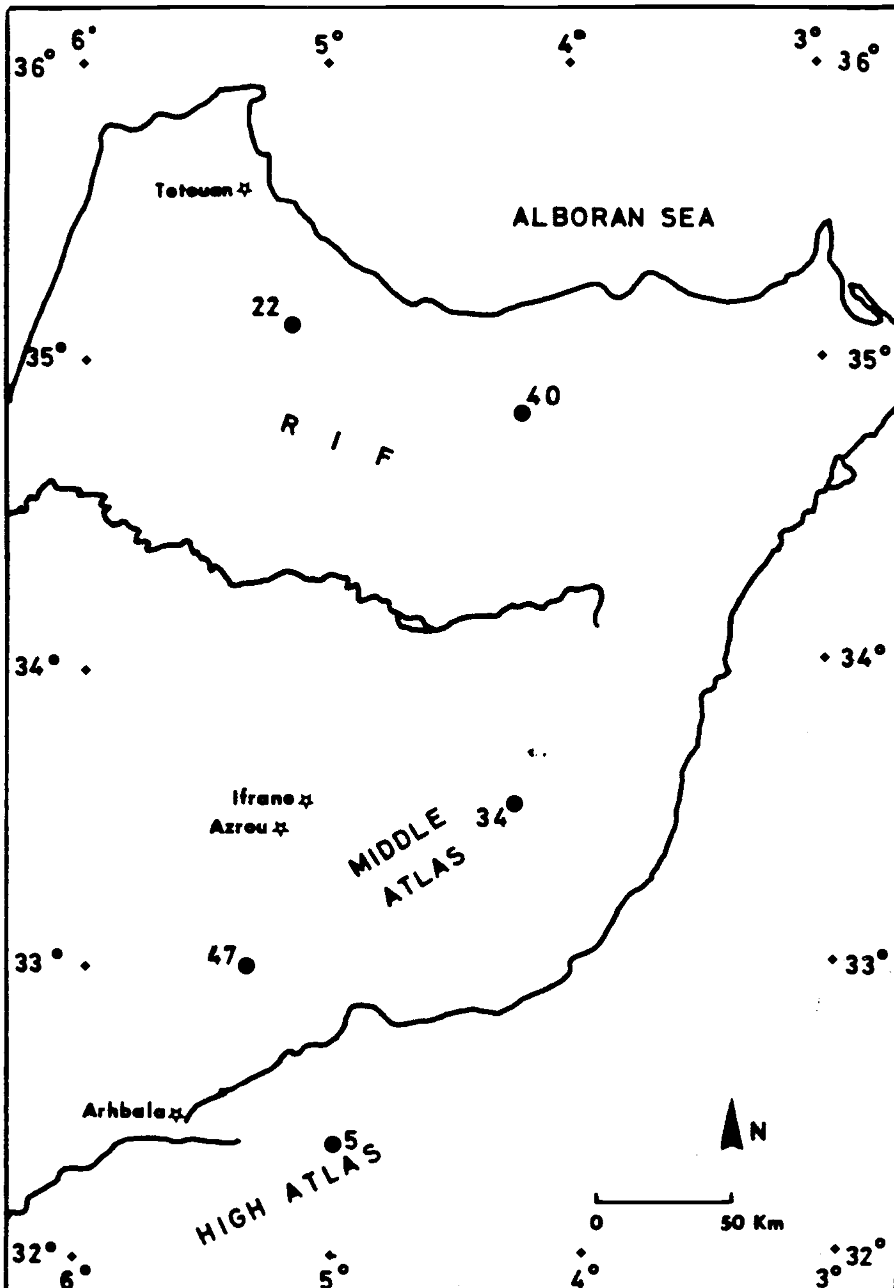

Figure 1. Location of tree-growth sites studied: 5 Afraskou (3050-Lv), 22 Jbel Lakrâa (3280-Lv), 34 Immouzer des Marmoucha (3420-Lv), 40 Tizi Ifri (3550-Lv), 47 Jbel Irhoud (3620-Lv).

$$X_{i.} = (\sum_{t=1860}^{1976} X_{it})/117$$

This centering of the data is performed in order to obtain a symmetrical projection of the variable-points in the principal components planes of the individuals in a way similar to the factorial correspondence analysis (FCA).

2. standardization of the tree-ring series (X rows):

$$X^s = (X^s_{it})$$

where:

$$X^s_{it} = X^c_{it}/S_i = (X_{it} - X_{i.})/S_i$$

$$S_i^2 = \sum_{t=1860}^{1976} (X_{it} - X_{i.})^2/116$$

Standardization of the tree-ring series scales their variance so they are more homogeneous than is normally the case.

The choice of these transformations instead of log transformation or first-differencing is justified as follows. Transformations should be applied to all tree-ring width series to achieve comparable results. This is the case for centering and standardization as defined above because of, respectively, the similarity with FCA and the heterogeneity of variances among tree-ring series. Log transformation is particularly recommended for removing exponential trends, but such trends are not present in the ring-width series of the old-adult stage trees selected for the present analysis. First-differencing, on the other hand, reduces first-order autocorrelation in the presence of trend but introduces autocorrelation in the case of absence of dependence over time or white noise. The autocorrelation of certain tree-ring series (sites 22 and 40 in particular) being much stronger than in the others, the effect of first-differencing would not be constant when applied to all sites. Both log transformation and first-differencing reduce the size component of tree-ring series, on which the definition of atypical years is based. As we observe below, the centering and standardization of the years never facilitate the interpretation of results, and the best graphical representation of atypical years provided by PCA is obtained with raw data.

Construction of the Temporal Walk

The PCA is carried out on the correlation matrix of either the raw or the transformed tree-ring series. Thus, yearly tree-ring widths are considered as the variables and trees or sites as the individuals in the PCA. The correlation matrix instead of the covariance matrix is used to perform PCA in order to avoid possible heterogeneity through time of the yearly tree-ring widths although they are variables of the same nature and defined on the same experimental unit.

The analysis is first carried out on all sites together by using the master chronologies and then by using all individual chronologies together. The analysis is also carried out on the individual chronologies for each site. When PCA is performed on master or individual chronologies of the 5 sites together, atypical years eventually detected concern the whole set of sites. Applied site by site, PCA detects atypical years for individual sites. The presence or absence of an individual in PCA sense -- *i.e.* a tree or a site -- with wide rings may influence the outcome of the analysis. However, in the case of a homogeneous set of individual tree-ring series, results are essentially the same whether PCA is carried out on the whole set or on a representative subset of the chronologies. This remark is valid for the data arrays under study.

By the correlation structure of a data array, we mean the structure of the corresponding correlation matrix in terms of its eigenvalues. Considering the individual chronologies of all sites, it must be pointed out that the between-sites correlation structure is altered by the within-site correlations, *i.e.* by the correlations between individual chronologies of the same site. This alteration will have repercussions in PCA performed on the variables, *i.e.* the yearly tree-ring

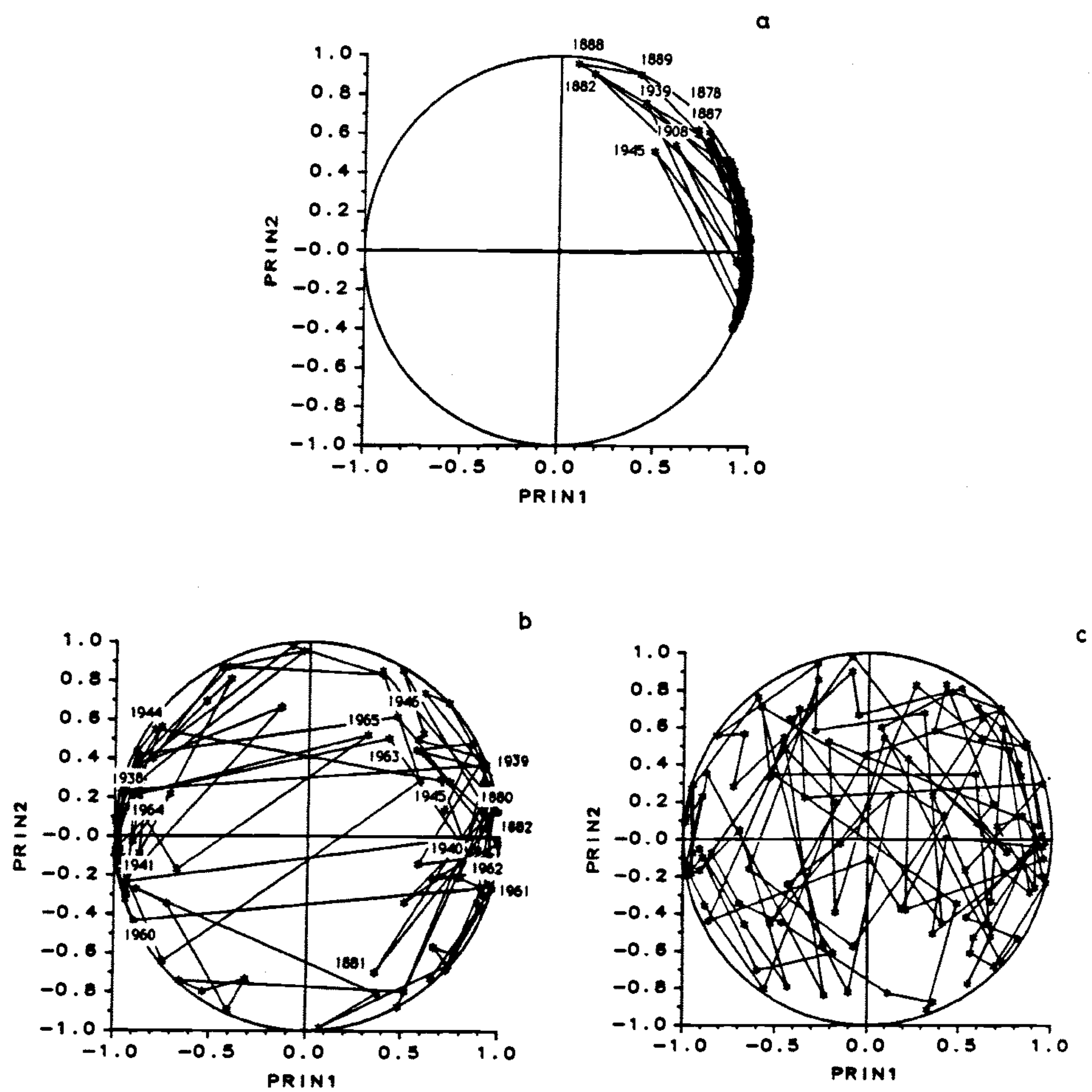

Figure 2. Principal components analysis performed on master chronologies: projection of the years in the plane of the two first principal components, PRIN1 and PRIN2, for raw data (a), centered data (b), and standardized data (c). The temporal walk is obtained by joining the years in chronological order in the plane supported by PRIN1 and PRIN2.

widths. In fact, it must be remembered that, from a mathematical viewpoint, all the positive eigenvalues relative to the variables in PCA differ only by a multiplicative constant from those relative to the individuals. This perturbing effect does not appear when master chronologies of all sites together are considered or, site by site, when individual chronologies are considered in sufficient numbers to reduce the influence of a single individual.

The construction of the *temporal walk*, which permits the detection of atypical years, is carried out by joining the variable-points (years) in chronological order in the plane supported by the two first principal components of the individuals (trees or sites). As a time series represents a temporal walk in the space of observations, the temporal walk provided by PCA is built in the space of

Table 1. Principal components analysis of sites 5, 22, 34, 40, and 47.

Chronologies	Eigenvalue (%)			Cumulative
	λ_1	λ_2	λ_3	Proportion (%)
Master 5 sites				
raw data	86.13	9.90	3.03	99.06
centered data	64.64	20.80	10.18	95.62
standardized data	39.82	25.33	18.04	83.19
Individual 5 sites				
raw data	62.51	9.62	6.67	78.89
centered data	30.78	17.15	13.05	60.98
standardized data	17.22	13.11	11.55	41.88
Individual site 5				
raw data	61.41	15.67	11.99	89.07
centered data	37.35	26.03	19.82	83.23
standardized data	34.50	26.60	21.08	82.18
Individual site 22				
raw data	55.18	18.90	11.84	85.92
centered data	37.71	23.44	16.16	77.31
standardized data	37.88	20.79	18.11	76.78
Individual site 34				
raw data	74.87	11.63	8.53	95.03
entered data	33.59	30.46	20.22	84.27
standardized data	34.01	26.12	20.87	81.27
Individual site 40				
raw data	35.04	25.45	20.02	80.51
centered data	35.85	25.40	16.30	77.55
standardized data	31.83	25.63	19.67	77.13
Individual site 47				
raw data	48.38	26.60	13.57	88.55
centered data	42.03	25.04	15.84	82.91
standardized data	33.43	30.53	15.33	79.29

principal components. The latter temporal walk, as the former, may be characterized by a strong first-order autocorrelation, which implies that a year can sometimes be detected as atypical only because the previous one is. This must be taken into account in the interpretation.

Table 2. Atypical years detected by principal components analysis carried out on raw data.

(1)	(2)	(3)	(4)	(5)	(6)	(7)
1878	1882	1867	1860	1880	1878	1860
1882	1888	1868	1868	1882	1880	1861
1887	1889	1879	1869	1883	1882	1862
1888	1890	1888	1882	1884	1932	1863
1889	1908	1893	1885	1886	1933	1864
1908	1939	1894	1886	1887	1934	1865
1939	1945	1898	1887	1889	1935	1866
1945		1899	1890	1904	1936	1867
		1900	1893	1905	1937	1868
		1901	1905	1906	1938	1869
		1905	1907	1907	1939	1870
		1906	1917	1908	1940	1871
		1908	1924	1910	1945	1872
		1929	1965	1943	1950	1873
		1934	1967	1945	1955	1875
		1939		1946	1959	1876
		1941		1966	1960	1879
		1944			1968	1884
		1946			1969	1885
		1947				1886
		1968				1918
		1976				1919
						1923
						1924
						1925
						1926
						1941
						1942

1. Master chronologies of all sites.
2. Individual chronologies of all sites.
3. Individual chronologies of site 5.
4. Individual chronologies of site 22.
5. Individual chronologies of site 34.
6. Individual chronologies of site 40.
7. Individual chronologies of site 47.

RESULTS AND DISCUSSION

The interpretation of the PCA is carried out on the basis of Table 1 and Figures 2 and 3. Table 1 describes the numerical results of the PCA. The eigenvalues (%) represent the proportion of the dispersion explained by the corresponding principal components. Those figures represent the temporal walk provided by PCA performed on the correlation matrix of either the raw or the transformed tree-ring data.

For raw data, the first principal component is easy to interpret for the master chronologies (Figure 2a) and, to a lesser degree, for the individual chronologies (Figure 3a). In fact, all scores being positive and almost equal, the first principal component is obviously a size component,

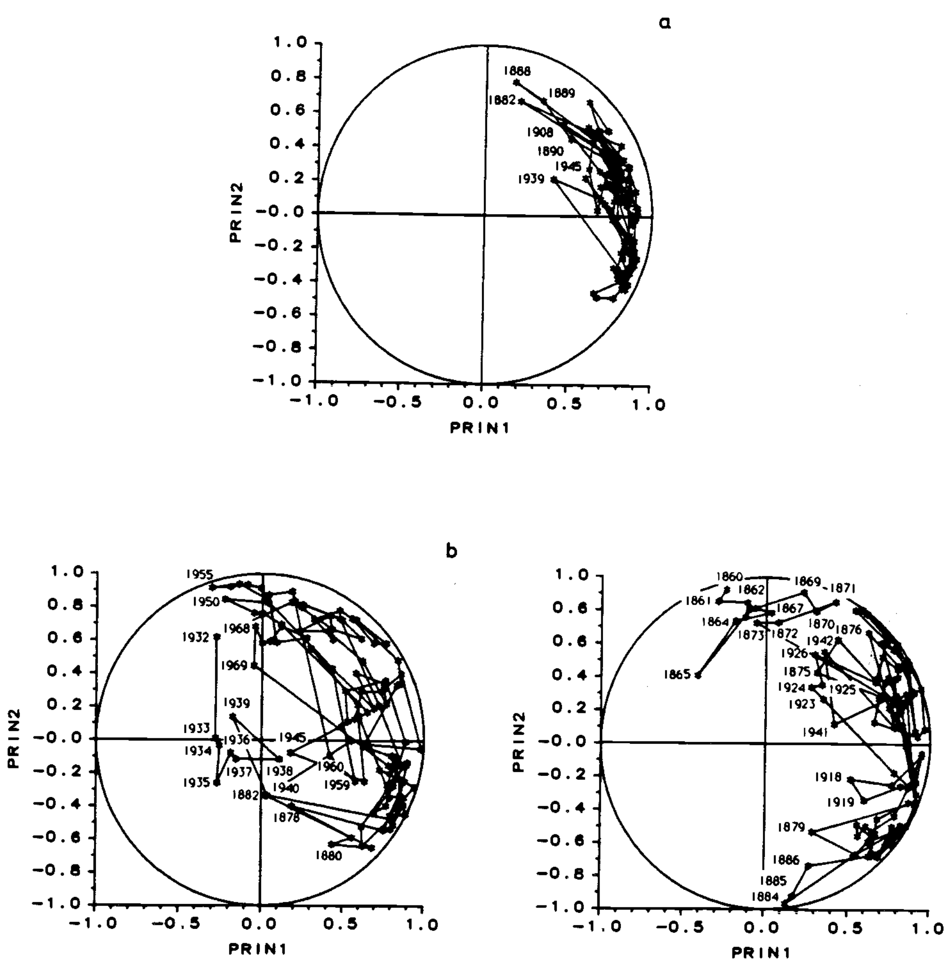

Figure 3. Principal components analysis performed on individual chronologies: temporal walk and atypical years are represented for all sites (a) and for two sites in particular, 40 (b) and 47 (c).

ranging from 86% of the dispersion for the master chronologies to 35% of the within-site dispersion for site 40 (Table 1). Low values of yearly ring-width projections onto this size component correspond to years of anomalous growth, *i.e.* to atypical years as defined in the introduction. The temporal walk allows the detection of these years because they are out of the sequence of normal years in the principal components planes of the individuals. The second principal component is supported by atypical years for raw master chronologies (Figure 2a). The atypical years are detailed in Table 2 for each site and for all sites together.

Most of the atypical years detected for the five sites on the basis of their master chronologies correspond to narrow rings for sites 5 and 40, to wide rings for site 22, and to mixed ones for sites 34 and 47 (Figure 4). The repartition of narrow and wide tree rings at atypical years has to be linked to the sensitivity of the five sites. The detection of atypical years for each site in particular is more difficult through the correlation structure among individuals, which reflects the high level of noise in the system. This fact is particularly verified for site 40 (Figure 3b), whose individual chronologies vary in their major part in a nonconcordant manner, reflecting a hardly satisfactory synchronization. This problem does not exist when PCA is performed on master chronologies (Figure 2a) for which the parallelism is usually broken only at atypical years. The high number of atypical years detected for site 47 (Figure 3c) may be attributed to age differences among trees and the autocorrelation of the individual chronologies. In fact, their nonconcordance begins at their first year (1860) and extends up to 1873.

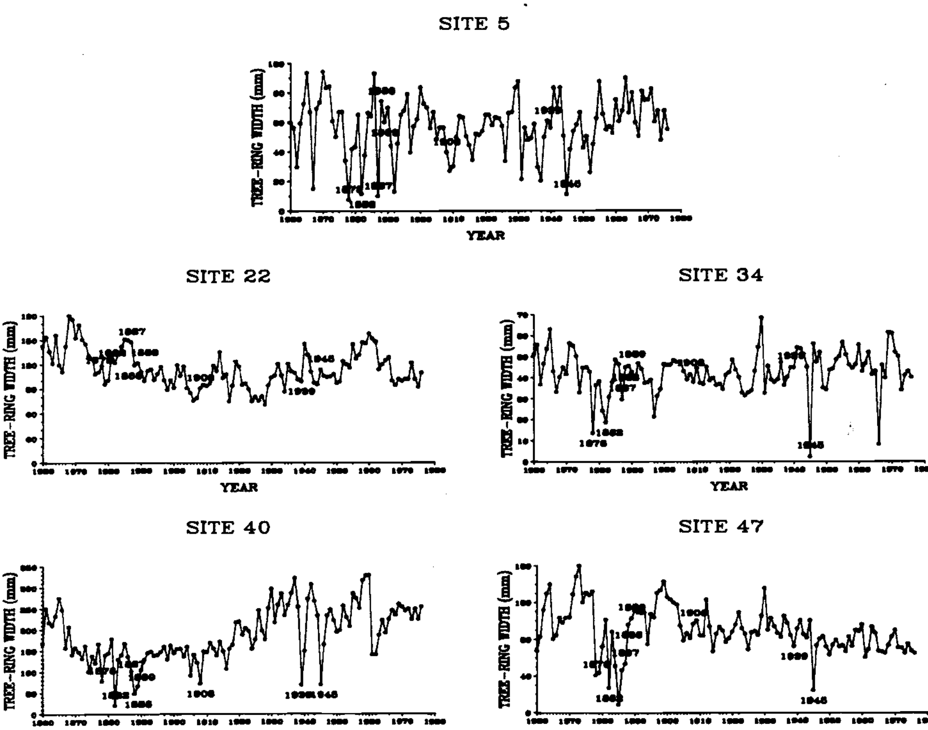

Figure 4. Master chronologies of sites 5, 22, 34, 40 and 47: they evolve in a quasiparallel way and differ by a constant, except at atypical years.

The first principal component on the centered data (Figure 2b) is a sort of size component in which an atypical year is characterized by a right-left-right movement or its reciprocal. A nonimmediate return in this movement may be attributed to the strong autocorrelation of the chronologies.

Interpretation of the standardized data (Figure 3c) is impossible, although the third principal component takes important values for all atypical years (Dutilleul and Till 1988).

CONCLUSIONS

PCA is a powerful method to synthesize data arrays of time series and of tree-ring series in particular. By the construction of a temporal walk in the plane supported by the two first principal components, PCA of tree-ring series provides a pertinent graphical representation with easy detection of atypical years for raw data when master chronologies are considered. PCA may be performed on individual chronologies for each site only if a sufficient number of individuals is available. The number acceptable depends on the length of the chronologies, five individuals appearing to be a minimum in our case. The atypical years obtained in this analysis are specific to the five cedar forests studied. Ecological interpretation will be possible only if the method is applied to the whole set of chronologies available for Morocco.

ACKNOWLEDGMENTS

The authors are particularly indebted to the Laboratoire de Palynologie et de Dendrochronologie (Université Catholique de Louvain) for permission to use the data analysed in this paper. The authors gratefully acknowledge Professor G. Gérard (Unité de Biométrie et Analyse des Données, Université Catholique de Louvain) for fruitful discussions of the methodological aspects of this work. Many thanks also to Dr. E. Le Boulengé (Unité d'Ecologie et de Biogéographie, Université Catholique de Louvain) for improving the English editing and to the anonymous referees for their valuable remarks.

REFERENCES

Dutilleul, P., and C. Till
1988 Principal components analysis and discriminant analysis in dendrochronology. In *Wood and Archaeology (Bois et Archéologie)*, edited by T. Hackens, A. V. Munaut, and C. Till, Pact 22:37-52.

Fritts, H. C.
1976 *Tree rings and climate*. Academic Press, London.

Guiot, J.
1981 Analyse mathématique de données géophysiques. Applications à la dendroclimatologie. Doctoral Thesis, Catholic University of Louvain, Louvain-la-Neuve.

Guiot, J., A. L. Berger, A. V. Munaut, and C. Till
1982 Some new mathematical procedures in dendroclimatology, with examples from Switzerland and Morocco. *Tree-Ring Bulletin* 42:33-48.

Jacoby, G. C., and E. R. Cook
1981 Past temperature variations inferred from a 400-year tree-ring chronology from Yukon territory, Canada. *Journal of Arctic and Alpine Research* 13:409-418.

Lebart, L., and J. P. Fénelon
1975 *Statistique et informatique appliquées*. Dunod, Paris.

Lebart, L., A. Morineau, and J. P. Fénelon
1979 *Traitement des données statistiques: méthodes et programmes*. Dunod, Paris.

Lebart, L., A. Morineau, and N. Tabard
1977 *Techniques de la description statistique: méthodes et logiciels pour l'analyse des grands tableaux*. Dunod, Paris.

Peters, K., G. C. Jacoby, and E. R. Cook
1981 Principal components analysis of tree-ring sites. *Tree-Ring Bulletin* 41:1-19.

Till, C.
1985 Recherches dendrochronologiques sur le Cèdre de l'Atlas (*Cedrus atlantica* (Endl.) Carrière) au Maroc. Doctoral Thesis, Catholic University of Louvain, Louvain-la-Neuve.

1987a Caractéristiques des chronologies de cernes définies sur le Cèdre de l'Atlas (*Cedrus atlantica* (Endl.) Carrière) au Maroc pour étudier l'évolution des forêts et du climat. *Dendrochronologia* 5:143-181.

1987b The summary response function of *Cedrus atlantica* (Endl.) Carrière in Morocco. *Tree-Ring Bulletin* 47:23-36.