Annals of Silvicultural Research
45(1), 2020: 62-75
https://journals-crea.4science.it/index.php/asr

Technical note

Collection: "4th Italian National Congress of Silviculture"
Torino, 5-9 November 2018

Hierarchical Factor Classification of Dendrochronological Time-Series

Sergio Camiz[1*], Francesco Spada[2], Jean-Jacques Denimal[3], Sergio Piraino[4]

Received: 29/10/2019 Accepted: 15/01/2020 Published: 09/09/2020

ABSTRACT In this paper, Hierarchical Factor Classification (HFC), an exploratory method of classification of characters is introduced, in comparison with Principal Component Analysis (PCA) in order to show its advantages, in particular when dealing with time series. Exploratory data analysis may play a very relevant role in the understanding of the structure of a data set prior the use of statistical methods – as hypothesis testing and inference, and models. The study of tree-rings time series through exploratory methods may also take advantages, by allowing some interpretation to be further checked via a small number of statistical tests. In particular, while providing overall results close to those of PCA, HFC complements it, by providing a classification of the time-series and estimating a representative chronology for each group, common to the clustered ones. As case study, a data set is taken from literature, composed by five synchronous 79 years-long chronologies of *Pinus pinea* L., from five different populations scattered along the Tyrrhenian coast in peninsular Italy. HFC suggests how conveniently aggregate the chronologies, by showing similarities and differences between them, otherwise unnoticed, suggesting to limit the aggregation to three chronologies only.

KEYWORDS: chronologies, *Pinus pinea* L., Principal Component Analysis, Hierarchical Factor Classification, exploratory data analysis.

Introduction

This paper aims to introduce an exploratory method of classification of characters, Hierarchical Factor Classification (in the following, HFC: Denimal 2007), through a case study, and to show how its results may be interpreted and give way to further investigation. The method is able to deal with synchronous time-series – such as those used in dendrochronology – issuing a hierarchy in which each node (i.e. a formed group) is described through its factorial structure.

In particular, it may be used to ascertain to what extent a set of such time-series may be synthesized by one or more general chronologies, corresponding to a representative chronology naturally associated to the group. Based on a rationale analogous to Principal Component Analysis (in the following, PCA: Lebart et al. 2006), it has the advantage to build a hierarchy of the characters based on correlation and to allow consequent partitions – something that, based on PCA only, may not be done – while keeping both the factorial graphical representations of characters and units similar to those issued from PCA and the related interpretation aids.

The construction of a chronology – a site-level representation of tree growth (Speer 2010) – based on tree-ring width, late-wood density, or other characteristics of timber, is the basis of dendrochronology, aiming at dating specimens, artefacts, but also at estimating past climate, since direct measurements are missing beyond the length of the instrumental records and the width of the tree-rings is widely known to mirror climatic fluctuations (Cook and Kairiukstis 1992, Rohli and Vega 2018). The fundamental assumption in dendroclimatology is that a climatic signal may be hidden into the growth of tree-rings and it is usually estimated with the mean of several synchronous tree-ring width time-series (Fritts 1976 and 2012, Boreux et al. 2009). In order to detect such signal and to obtain a good reconstruction, dendrochronologists must take crucial decisions about the tree species, the region of interest, and the sampling procedure (Cook and Kairiukstis 1992, Saint George et al. 2008). Therefore, they must rely to an accurate data analysis of the collected data to achieve their task, in particular through a correct use of both data analysis and statistical methods to deal with such matters.

From the data analysis point of view, several different methods are required to achieve this task: let us briefly quote the identification of common signals in a set of synchronous time-series on one side and the relations between tree-growth and climate on the other, which need different tools to be carried out. In this paper we concentrate on the first step, namely the search for a common signal, which is the key item to be used in identifying past climate conditions based on tree-rings chronologies: as the study of relations between them requires other specific exploratory tools, they had better discussed separately.

In dendrochronology, multidimensional data analysis techniques, in particular PCA, have been

1 ISPC - CNR (National Research Council) - Roma - Italia
2 Dipartimento di Biologia Ambientale, Sapienza Università di Roma - Italia
3 Université des Sciences et Technologies de Lille - France
4 IANIGLA-CONICET Mendoza - Argentina
*Corresponding author: sergio@camiz.net

http://dx.doi.org/10.12899/asr-1968

largely used since long. PCA was first used in multiple regression context to prevent too strong correlations between regressors (Fritts et al. 1970, Fritts 1976, Fowler 1988, Briffa et al. 2001, Patskoski et al. 2015) and to synthesize the climatic data, which should be matched with the chronologies (LaMarche and Fritts 1971, Biondi et al. 2001, Gray et al. 2004). Both LaMarche (1974) and Peters et al. (1981) discuss in detail the use of PCA for the identification of chronologies. In parallel, different classification techniques have been used on the same data sets to identify groups of tree-ring time-series susceptible to be grouped into distinct chronologies (Piovesan et al. 2005, Mazza et al. 2014, Touchan et al. 2016) according to several purposes and criteria. Litton and Zainodin (1991) propose a complex model to choose among regional and national chronologies. More recently, PCA was sometimes applied to both site chronologies and tree-ring time-series to assess, among other issues, their correlation and the uni-dimensionality of a set of series at hand (Papadopoulos et al. 2009, Bunn et al. 2013, Papadopoulos 2016): a task which HFC may deal more appropriately, given the concurrent issue of groups formation and their factorial structure, which provides in addition an overall PCA-like study..

Thus, PCA is a traditional exploratory technique for this kind of data. By exploratory (*sensu* Tukey 1977), we mean those methods useful to study data without any a priori statistical model, which "let the data speak for themselves" (Benzécri 1973) and which help to identify relations and structures, to be further tested – on other samples, taken for the purpose – through statistical methods, to be confirmed. Nevertheless, as in many other frameworks, the use of exploratory data analysis techniques has often been misunderstood or misused, due to the illusion that they could be handled as statistical or modelling methods. The most relevant misuses of PCA are the choice of a very limited number of principal components without any consistent statistical reason (which, incidentally, are far from being reliable, Camiz and Pillar 2018); their use to model the whole data set, without checking the non-reconstructed part, which should observe the ordinary residuals' conditions: randomness, independence, and equal variance; the use of the first principal component as a common signal, without checking its correlation with the original time-series, etc. Indeed, the exploratory use of these techniques should be limited to exploration, namely to facilitate the researchers to study their own data in the most appropriate way, from the most evident results to the most hidden ones. In this sense, this way of dealing with them is really limited or ignored outside the data analysts community. On the opposite, either they are confined to very rough circumstantial inferences extracted from their output, or their results are used as final statements, instead of taken as hypotheses to be tested.

Hence, also PCA had rather be limited to the mere study of the data, while taking into account its results in the whole subsequent study. On the opposite, most of the quoted works use PCA in an instrumental way – without a deep analysis of its results – and/or for a graphical synthetic representation of issues relevant for both methodology and study aims, but not really taking advantage from its use. In fact, it would help to better understand data structure and identify special patterns to be more accurately inspected, when necessary, with other statistical techniques. In the case of chronologies, PCA may reveal, in decreasing order of relevance, linearly uncorrelated time-series, whose weighed sum would approximate those in the data-set; they represent uncorrelated sources of variation – which ought to be determined –, either common to all original series or specific to only some, but they do not necessarily represent a common signal.

The alternative exploratory method proposed here, HFC, creates groups of time-series in a hierarchical structure and for each one produces a pair of orthogonal uncorrelated time-series. Of them, the first may be considered a factor common to all characters of the group and the second showing their differences. In addition, by checking the numerical results – in particular the correlations – one may ascertain the uni-dimensionality of each group, and, should all series in the group be positively correlated with the factor, this one may be considered an approximate estimation of a common signal. Thus, the method may be applied to dendrochronological studies, for the search of common chronologies in a more straightforward way than PCA, without losing the interpretability of PCA.

Indeed, HFC has been already used in dendrochronology by both Piraino et al. (2013) and Stafasani and Toromani (2015) in parallel with PCA: in the first paper no direct comparison of the results was reported, since the study was focused on the response of plants to climate, while in the second HFC helped in identifying two groups of sites, opposite on the first PCA factorial axis, showing its specific ability to classify what PCA only outlined. As both papers focused on specific issues, HFC was used after a short presentation only, referring to the literature for details. We believe that the method deserves a better presentation in the dendrochronology framework, to better illustrate its advantages, in particular with respect to PCA: it is what we are aiming at in this paper. For this reason, we are not dealing with new original data, but with those already studied by Piraino et al. (2013), concerning chronologies of *Pinus pinea* L. wood ring-width: the comparison might show the quality of HFC results. In particular, in this paper we try to understand whether they may be gathered in a general chronology or they had rather separated in different groups.

This paper is structured as follows: in " The chronologies" section, the case study is introduced; in "Principal Component Analysis" section the essentials of PCA are reminded, in order to better ground HFC rationale, which is presented in detail; then, in the "Hierarchical Factor Classification" section the results of both PCA and HFC applied to the *P. pinea* chronologies are reported, to put in evidence the issued differences and to show the HFC 's advantages. Discussion and conclusion will follow.

The chronologies

The 79-years long chronologies range from 1925 to 2003 and were built by Piraino in sites along the Tyrrhenian coasts of the Italian peninsula: San Rossore, Cecina, and Duna Feniglia in Tuscany, and Castelporziano and Circeo in Latium. These sites are aligned NW – SE and are located close to the coastline of the Tyrrhenian Sea, between 43°43' and 41°18' North. All the stands are artificial, but natural regeneration is intense. They grow under Mediterranean climatic conditions, locally characterized by summer drought, ranging from one to three months. In these sites the species grows on sandy soils. The pine populations of San Rossore, Cecina, and Circeo originate from plantations carried out during the first half of the 20th century, while the populations of Duna Feniglia are some decades older and at Castelporziano the pine stands date back to 18-19th century (see Piraino et al. 2013, for further details). In Figure 1, the patterns of the five chronologies along their 79-years long time-span are represented, the vertical order of the chronologies corresponding to their geographical one. Note that, to keep all chronologies comparable, they have been standardized, as it will be done for the results. This means that all have zero mean and unit variance and no physical unity of measure. Indeed, no loss of information occurs, since a simple transformation may reconstruct the original data.

The data analysis methods

From the mathematical point of view, a data table composed by a set of synchronous chronologies of tree-ring widths represents a multidimensional time-series, i.e. a matrix whose characters in column are time-series referring to either similar or different items, with the only constraints that the items are synchronous and observed at regular intervals of time. Thus, the units on the matrix rows correspond to each measurement time and they are naturally ordered accordingly.

In data analysis, the current idea of information of a data table, resulting from its intrinsic variation, is measured by its inertia, that is the weighed sum of its squared values. Given $\boldsymbol{X}$, a quantitative data table with n units by row and p characters by column, its inertia is

$$Inertia(X) = \sum_{i=1}^{n}\sum_{j=1}^{p} w_i x_{ij}^2$$

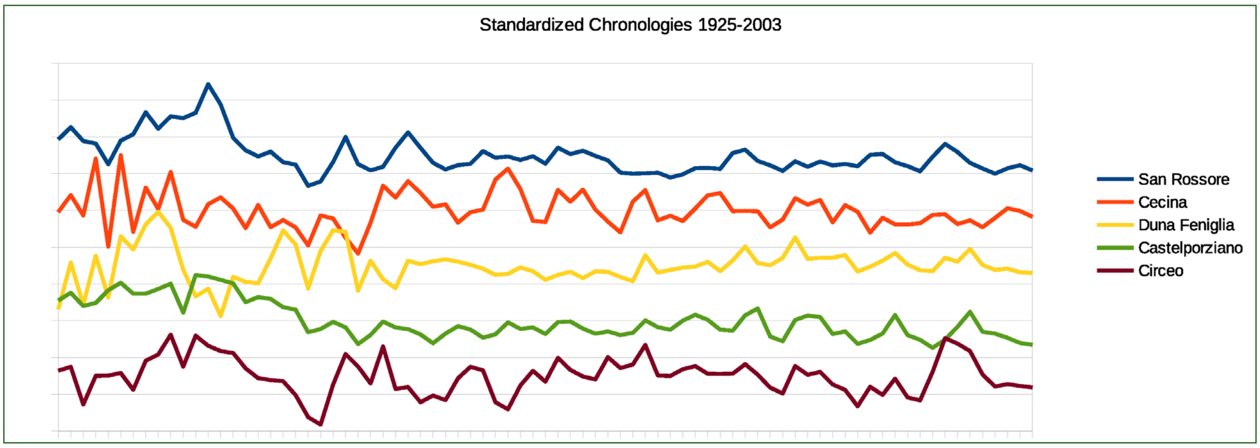
with w_i the weight given to the unit i such that

$$\sum_{i=1}^{n} w_i = 1$$

The weights usually are valued all $1/n$, but may be different should one wish to give different relevance to some units in the analyses. The inertia is a key concept in exploratory data analysis – because it is considered a measure of information – and most methods are based on its decomposition in independent components, through maximization.

Thus, a current practice prior the analyses is to prevent that the characters may bias the results, due to the different inertia caused trivially by the different units of measure. This is achieved by standardizing each character of the data table, i.e. to centre it around its mean and to reduce it to unit variance/inertia. This way, its inertia/information is worth 1, the same for each character, hence that of the total data table is worth p, the number of characters.

Figure 1 - The five standardized tree-ring widths chronologies of *P. pinea* of Central Italy under study. The top-down sequence mirrors the geographical NW-SE sequence of the site locations.



Unlike the classification of units, largely discussed in literature (see, e.g., Anderberg 1973, Gordon 1999), specific methods for the classification of characters have received minor attention so far. Lerman (1981) proposes a probabilistic hierarchical method, called Likelihood Linkage Analysis. The procedure VARCLUS of SAS package (Nelson 2001) is based on PCA and on the works of Anderberg (1973) and Harman (1976), aiming to define nearly unidimensional groups through a divisive algorithm. More recently, Vigneau et al. (2006) proposed an iterative non-hierarchical method based on the reallocation of characters with K-means (MacQueen 1967) style around the first principal component taken as centroid.

The alternative exploratory method of classification we are introducing in this paper, Denimal's Hierarchical Factor Classification (HFC, Denimal 2001 and 2007, Camiz and Denimal 2006, Camiz et al. 2006), is common to the two traditional families of data analysis methods: ordination and classification (Whittaker 1973). It proved to be consistent with classical hierarchical classification methods (Camiz and Pillar 2007) but in addition it provides a pair of principal components for each built group. For this reason, it may play the double role of classification and ordination technique, thus providing a deeper understanding of the relations between time-series. As such, HFC may be applied to dendrochronological studies, for the search of common chronologies, as it will be shown on the case study.

Principal Component Analysis

In order to better understand the rationale of HFC, we briefly remind here the essentials of PCA, referring to literature for technical details (see, e.g. Benzécri 1973, Bry 1994, Jolliffe 2002, Lebart et al. 2006, Husson et al. 2017). PCA aims to create a set of new independent characters in decreasing order of corresponding inertia, whose first ones synthesize at the best the information contained in the table; thus, the pattern of both units and original characters on 2-dimensional graphics, may be used, together with the corresponding numerical results, for a progressive study of a data table: it may be carried out step by step, from the easiest evident and most informative relations to the most hidden ones. The basic principle of PCA is to decompose the inertia of the data table in p uncorrelated principal components, i.e., the new characters, composed as weighed sums of the p original ones. To build them, PCA extracts the eigenvectors of either the covariance or the correlation matrices between the original characters. They are sorted according the decreasing inertia they are accounted for, measured by their corresponding eigenvalue. On planes spanned by pairs of principal components, the units may be optimally projected, minimizing the bias due to the projection and approaching at the best their original position. To principal components, principal axes are associated, defining planes on which the original characters may be projected as vectors within the so-called circle of correlations. Here, the cosine of their angle with the axes corresponds to their correlation: as they are maximized, they may be used to interpret the factors through the original characters. On these planes, additional units and characters may be projected as illustrative or supplementary, to show their relation with both the factors and the original characters. This way, one may check the ability of principal components – hence of the original characters – to approximate them as well as use them as an aid to interpret the principal components.

The use of PCA considering time-series as characters and observation time as units is feasible (Bry 1994), even if the observations are not independent. The principal components are time-series themselves, representing the evolution along time of the factors influencing the time-series table. On the opposite, it is not sure that they may be considered as principal component chronologies (PCC, *sensu* Peters et al. 1981), this being possible only for the first principal component, provided that all its coefficients are positive.

Hierarchical Factor Classification

Based on PCA results, no classification of characters seems possible, because it does not provide any measure of association between characters on which to ground a method, neither the observation of the circles of correlations issued by PCA, nor the computed correlations may help in this task, so that one has to apply to other concurrent methods.

A special attention deserves PCA of pairs of characters: unlike the principal components of a larger data table, whose explanation may sometimes be obscure – in particular for those following the first one – in this case it is easy to prove that the first principal component synthesizes what the two characters have in common – sometimes with opposite meaning – and the second which are their differences. This ease of interpretation led Denimal (2007) to develop HFC, aiming to combine in a single procedure the classification of characters with factorial methods analogous to PCA.

An ascendant hierarchical method of classification (Anderberg 1973, Gordon 1999, Lebart et al. 2006, Husson et al. 2017), builds a hierarchy on the objects at hand, i.e., a set of encapsulated partitions, and produces a dendrogram, namely a tree-graph, whose lower nodes are the objects, and the others are the step-by-step built groups, tied through arches downwards to the pair of joining groups and upwards to a group they form with another one; eventually, the last node corresponds to the whole set of objects forming one overall group. Thus, a partition is obtained by "cutting" the dendrogram at some suitable level. To build such a method, the following items are required:

(i) an association rule, to evaluate the similarity between the objects at hand: in HFC, association is measured by the second eigenvalue of a pairwise PCA;

(ii) an optimization method to be used to choose which objects/groups to join at each step: in HFC it consists in choosing the pair whose second eigenvalue is minimum;

(iii) a method for upgrading the associations at each step: in HFC the pairwise PCA is run between the representative character of the newly formed group and each representative character of all others;

(iv) should a partition be required, a criterion to choose where to cut the dendrogram is necessary, usually based on the hierarchy index, i.e. the optimum value found at each step: in HFC one may refer to the second eigenvalue, which measures the within group homogeneity.

As mentioned, HFC builds a hierarchy on a set of numerical standardized characters by computing PCAs on the covariance of pairs of characters. It operates as follows:

1. At the beginning each character is standardized and it is assumed to form a group by itself (a singleton), thus being also representative of it. Then the recursive algorithm is based on the following steps.

2. A pairwise comparison of the existing groups is done, by submitting to PCA their corresponding pair of representative characters (which, for the singleton are the original characters themselves), based on their 2×2 covariance matrix. Through PCA, the matrix inertia is split in two eigenvalues, the largest representing the information common to both representative characters, hence to all characters gathered in the group, and the second the one concerning their differences.

3. The pair of groups showing the minimum second eigenvalue issued by its PCA is selected as the one showing least differences between its components, hence being the most homogeneous.

4. The two groups of characters corresponding to the selected pair are merged in a new node of the hierarchy.

5. The first principal component of this PCA is chosen as representative component of the newly formed group.

6. The first eigenvalue, i.e. the inertia of the representative component, is a share of the total inertia common to the characters in the node.

7. The coefficients of the second principal component measure the distance of each character in the node to the representative component; hence this may be called the differences component.

8. Its corresponding second eigenvalue is chosen as the hierarchy index of this node.

9. Two graphical representations result, based on the two extracted components: *(i)* a circle of correlation, showing the correlations with them of the components of the two merged groups and of all characters belonging to the formed group, and *(ii)* a principal plane, showing the pattern of the units as seen by that same pair of components. Both planes are interpreted the same way as in PCA, allowing to appreciate the relations between characters as well as the corresponding pattern of the units.

If the characters are p, the steps from 2. to 9. are repeated $p - 1$ times, obtaining a complete hierarchical classification of the characters, together with pairs of principal components associated to the nodes. To define a partition, one may check through the nodes indexes the amount of differences between the characters of the group, considering that it has little sense to gather uncorrelated characters.

It is noteworthy to observe that the representative components may be projected on the ordinary PCA of the data table as illustrative elements. This allows an interoperability between the two methods and may contribute to a better understanding of the problem under study. Note that the representative components of different nodes need not to be uncorrelated, whereas the second usually exhibit low correlation both with the representative components of the other nodes and within themselves. Despite this, with the representative component of the first node, they provide a decomposition of the total inertia.

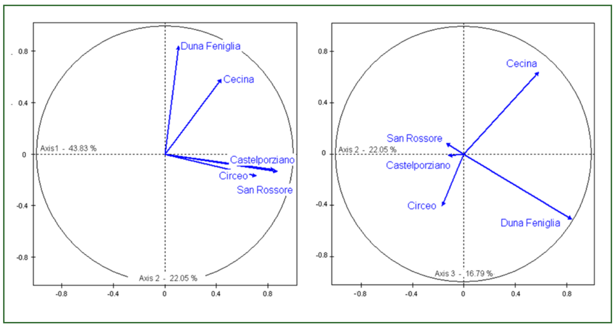
Dealing with time-series, in particular with chronologies, it results that the representative time-series of each group is situated within the directions of those forming the group. If its correlations with them have all the same sign, it may be taken as positive, so that, being a weighed average, it constitutes a kind of centroid of them. Should some be opposite in sign, the representative component is usually understood as the opposition between a dipole of characters, concurring to its interpretation with opposite meaning (see Denimal 2007, for the technical details). Dealing with chronologies, in the case of concordance of signs, the representative time-series may be adopted as representative chronology of the group. On the opposite, should a dipole result for some node, the group can be split in two opposed subgroups and two (opposed) common chronologies should be estimated in some other way. Given the non-orthogonality of the representative chronologies of different groups, their interpretation may only be based on the chronologies forming the group, but nevertheless they are better situated and interpretable than rotated and oblique principal components, sometimes preferred to classical PCA in dendrochronological studies (Büntgen et al. 2007, Frank and Esper 2005, Leland et al. 2013) even for classification purposes.

Results

PCA results

In Table 1, the eigenvalues issued by the PCA of the five chronologies are reported along with the percentage of inertia explained by the corresponding factors and the cumulate percentage. Here, we take into account the recommendation of Jolliffe (2002) – to consider relevant the dimensions whose inertia is at least 0.7 – hence we shall assume suitable to take into account three principal components, which summarize 82.67% of total inertia.

Figure 2 - Representation of the five tree-ring chronologies of *P. pinea* on the circles of correlations on the planes spanned by the factorial axes 1-2 (left) and 2-3 (right) issued by their PCA.



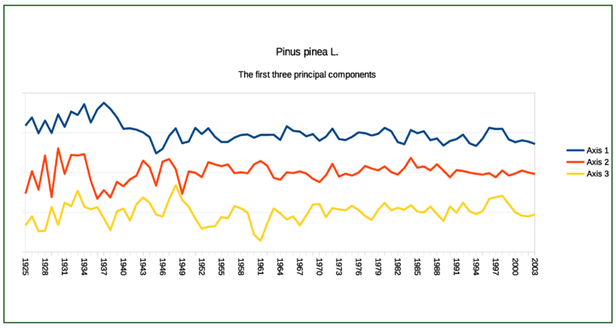
In Figure 2 the chronologies are represented on the circles of correlations on the planes spanned by the axes 1 and 2 (on the left) and by the axes 2 and 3 (on the right) of PCA, respectively. In the circle on the left, three chronologies, namely San Rossore, Castelporziano, and Circeo are oriented very close to the first axis, the latter less well represented than the other two. The remaining two chronologies are oriented towards the second axis, in particular Duna Feniglia, which appears nearly orthogonal to the said group of three. Cecina and Duna Feniglia result opposed along the third axis, as shown in the circle on the right. In this plane, both chronologies are pretty well represented and their nearly independence corresponds to the nearly right angle between them.

In Figure 3 the patterns along time of the three chronologies corresponding to the first three principal components are represented. By comparing them with the original chronologies shown in Figure 1, it is worth to point out the resemblance of the first one with the chronologies of the group of three and of the second one with Duna Feniglia.

Table 1 - The eigenvalues issued by PCA of the five chronologies of *P. pinea.* In the columns: the number, the eigenvalue, the percentage of total inertia attributed to the corresponding factorial axis, and the cumulate percentage of inertia.

Number	Eigen value	Inertia %	Cumulate %
1	2.191	43.827	43.827
2	1.103	22.054	65.881
3	0.839	16.786	82.667
4	0.557	11.143	93.810
5	0.309	6.190	100.000

Figure 3 - The pattern along time of the first three factors issued by the PCA of the five tree-ring chronologies of *P. pinea*.



In Table 2 the correlations between the chronologies and the five issued principal components are reported: just as an indication of relevance, consider that, should two characters be random and independent, the p-value associated to 5% level of significance of its appropriate statistics (a student t with n-2 degrees of freedom, Kendall and Stewart 1973) for n=79 observation is 0.222. Indeed, this is not true for correlations either between factors and variables or between time-series; thus, we took it only as a threshold for the correlations to consider in the discussion and nothing more: in the tables they are shown in boldface. in agreement with Figure 3, with the first principal component they are high for San Rossore, Castelporziano, and Circeo and medium for Cecina; with the second one are high high for Duna Feniglia and medium for Cecina; and with the third one medium and opposed for Duna Feniglia and Cecina.

Table 2 - Correlations between the five chronologies of *P. pinea* and the principal components issued by their PCA. Here, the p-value for the correlations significance at 5% level is 0.22. The correlations significant at this level are in bold.

	Axis 1	Axis 2	Axis 3	Axis 4	Axis 5
San Rossore	**0.873**	-0.137	-0.084	-0.199	**0.415**
Cecina	**0.436**	**0.581**	**-0.638**	**0.249**	-0.057
Duna Feniglia	0.109	**0.838**	**0.511**	-0.148	0.044
Castelporziano	**0.850**	-0.123	0.010	**-0.363**	**-0.361**
Circeo	**0.710**	-0.169	**0.404**	**0.549**	-0.050

In Figure 4 the pattern of the years on the plane spanned by the first two principal components is shown. Note in particular the positions very far from the centroid of the years in the period 1925-1950, in which mayor variations occurred for all chronologies, with an opposite behaviour of Duna Feniglia and Cecina in the corresponding years.

One may attempt an interpretation, stating that a common signal represented by the first principal component – but weak for Duna Feniglia – is likely, while observing relevant differences between this one and Cecina and with the others.

Figure 4 - The pattern of the years on the plane spanned by the first two principal components issued by the PCA of the five tree-ring chronologies of *P. pinea*.

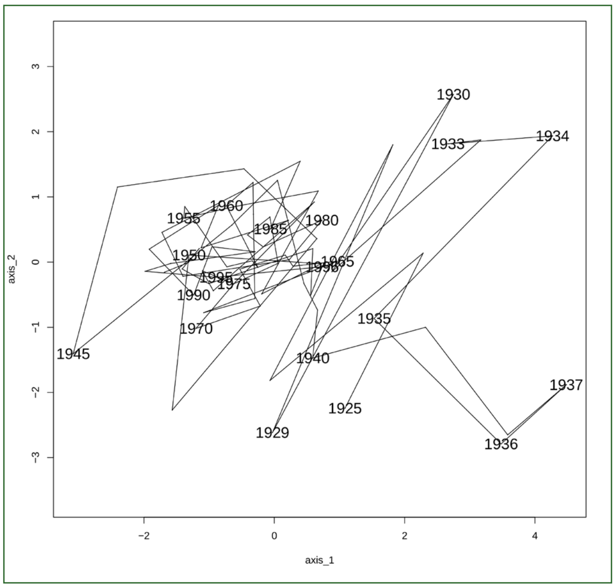


Figure 5 - The dendrogram resulting from HFC of the five tree-ring chronologies of *P. pinea.*

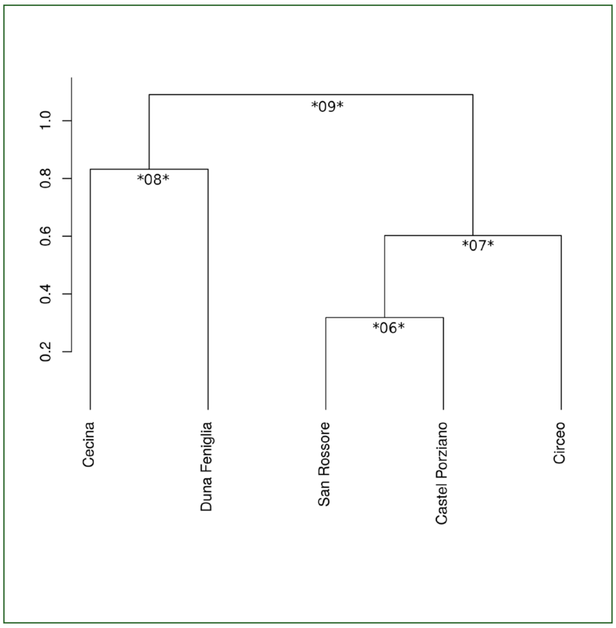


Table 3 - Results of the construction of the hierarchy on the five chronologies of P. pinea through HFC. In the columns: the number of the node, the number of groups of the corresponding partition, the two nodes merged at that level, the number of chronologies in the node, the fusion index, i.e. the second eigenvalue of the PCA of the two merged nodes, the percentage of total inertia attributed to its corresponding differences component, the cumulate inertia of all the differences components up to that node, the inertia attributed to the representative component and its share in respect to the corresponding PCA. At the end, the sum of the fusion indexes and the first eigenvalue of the upper node, with their attributed inertia, summing up to 100%.

Node	Groups	1st gr	2nd gr	Number	Fusion index	Inertia %	Cumulate inertia	1st axis %	Partial %
6	4	1	4	2	0.319	6.371	6.371	1.681	84.074
7	3	6	5	3	0.602	12.046	18.416	2.079	69.306
8	2	2	3	2	0.831	16.629	35.045	1.169	58.427
9	1	7	8	5	1.090	21.805	56.850	2.157	43.150
Sum of the hierarchy indexes					2.843		56.850		
First representative chronology					2.157	43.150			

HFC results

The dendrogram built by the HFC is represented in Figure 5: its topology in Newick format - that is by enclosing in parentheses the two groups merging at each level - is ((1,4),5),(2,3)).

Looking at the figure, a doubt may raise concerning the appropriate partition, since the thumbnail rule to search for the largest branches does not give evidence of a better one. Thus, we had rather inspect Table 3 where the numerical results concerning the hierarchy building are reported.

There, along with the sequence of fusion levels issued by HFC of the five chronologies, all the needed information may be found in the columns: the node number, the quantity of groups in the corresponding partition, the two nodes merged at that level, and the number of chronologies grouped together. Then, the fusion index, i.e. the 2nd eigenvalue of the PCA of the two merged nodes, the percentage of total inertia attributed to the corresponding direction, the cumulate inertia of all the 2nd principal components up to that node, the inertia attributed to the 1st representative character of the current node and its share in respect to the corresponding PCA are reported.

Note that the total data table inertia, worth 5, is partitioned according to the inertia along the first axis of the last node, a measure of the signal common to all chronologies, and the sum of the hierarchy indexes, an overall measure of their differences, both reported in the last rows of the table.

The inspection of Table 3 allows to select an appropriate partition, based on the differences between merging chronologies indicated by the fusion index given by the second eigenvalues: the smaller it is, the better is the ability of the node's representative time-series to synthesize those forming the group. As we may ground our choice on the principle that the formed groups must be uni-dimensional, we may apply the same Jolliffe (2002) recommendation upside down, i.e. by merging groups until the second eigenvalue does not exceed 0.7, since otherwise a second dimension within the group could be too relevant; indeed, it would not make much sense to aggregate in one group characters uncorrelated to each other. Based on this threshold, three groups may be considered, whereas the second eigenvalues of the last nodes, larger than this threshold, may suggest the existence of more than one time-series necessary to summarize all these group's chronologies. In our case, the group formed by San Rossore, Castel Porziano, and Circeo results, while Duna Feniglia and Cecina remain isolated.

In the following, the principal components issued by the HFA of each node are labeled by joining the node's number with A the first, i.e. the representative chronology, and with B the second one, the component of differences. In Figure 6 are represented the circles of correlations corresponding to the two last nodes *9* (to the left) and *8* (to the right). On them are represented the chronologies belonging to the node and the principal components of the merging groups. These do not appear in the circle to the right, because only two singleton chronologies are aggregated there. Looking at both, it is evident that in both cases the two merging groups of chronologies, those around their corresponding representative component in red to the left and the singleton Cecina and Duna Feniglia to the right, are little or no correlated, since the corresponding angle between the representative components is nearly squared. In fact, dealing with a 2 characters PCA, these are true angles and not projections. This justifies our choice to consider three groups.

In Figure 7 the patterns of the principal chronologies associated to the four nodes of the hierarchy are shown: it results that *6A*, *7A*, and *9A* are very similar to each other. This may be better appreciated looking at their correlations, reported in Table 4 where they range between 0.935 and 0.981, meaning that they are nearly the same; on the opposite, *8A* is quite different from the others, exhibiting a negligible correlation with *6A* and *7A* and a small one with *9A*.

Figure 6 - Representation of the five tree-ring chronologies of *P. pinea* (in blue) on the circles of correlations on the planes spanned by the two principal components of the nodes *9* (left) and *8* (right) issued by their HFC. In red the representative variables of the merging groups.

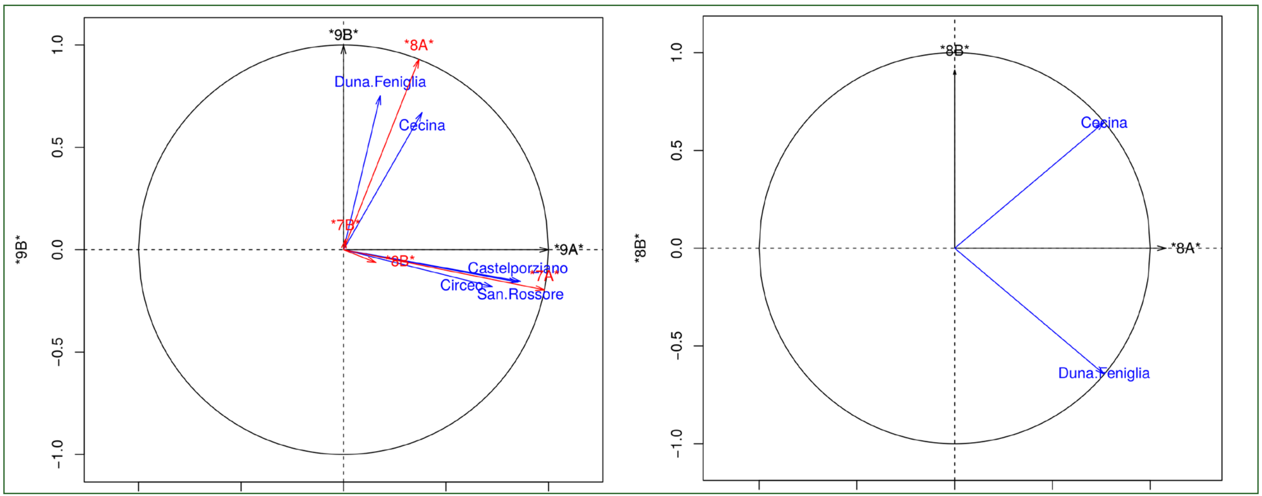
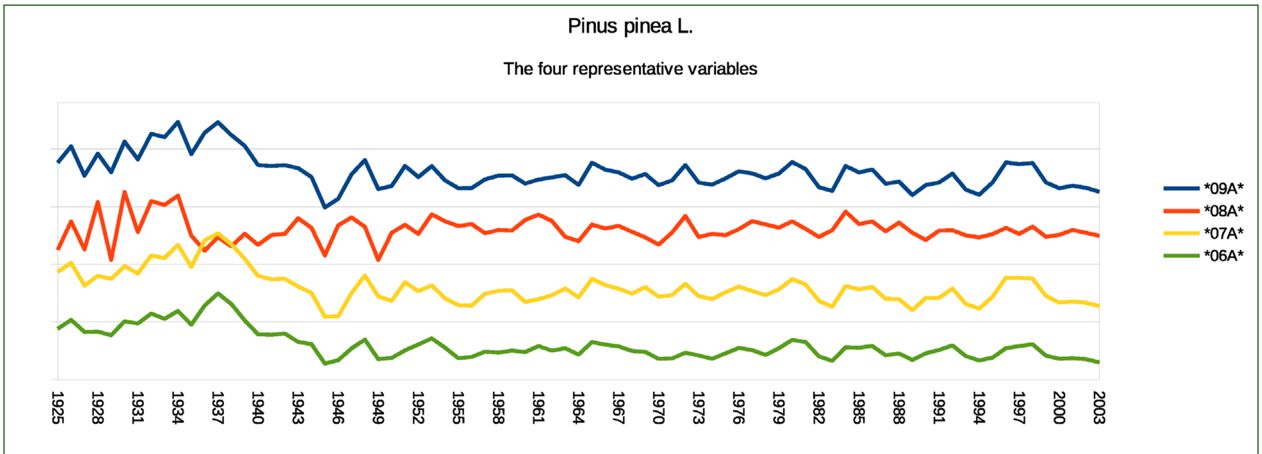


Figure 7 - The pattern along time of the representative chronologies of the highest four nodes issued by HFC of the five tree-ring chronologies of *P. pinea.*



This means that the chronologies grouped by node *8*, Cecina and Duna Feniglia, show a common time pattern independent from the others. Moreover, *8A* is highly correlated with *9B*, meaning that the difference between the two last nodes of the hierarchy is essentially due to the merging of the nodes *7* and *8*, which exhibit poor or no correlation between their chronologies.

In Table 5 the correlations between the five chronologies and the principal components of the nodes are reported. It is evident that the chronologies exhibit higher correlation with the representative one of the groups to which they belong. Thus, are evident the strong correlations of San Rossore and Castelporziano with *6A*, *7A*, and *9A*, of Circeo with both *7A*, and *9A*. Relatively high is the correlation of both Duna Feniglia and Cecina with both principal components of *8* and that of Circeo with both of *7* : this represents an analogous phenomenon, i.e. the non-relevant correlation between the two characters merging in that node, which, as said, gets it questionable. Note also the relevant correlation of both Duna Feniglia and Cecina with *9B*, the differences component of the last node, which confirms the scarce interest to gather them into a common overall chronology.

In Figure 8 the patterns of the years on the principal planes corresponding to nodes *9* (above) and *8* (below) are represented. The reading is easy: in the upper graphic, corresponding to node *9*, the years until 1949 are further from the origin, showing their larger variation; highest values of the group of three chronologies are situated on the right, with an opposition between the years 1930, 1932-1934 and 1936-1938 due to the high and low values, respecti-

Figure 8 - The pattern of the years on the plane spanned by the principal components of the nodes *9* (above) and *8* (below) of the hierarchy issued by the HFC of the five tree-ring chronologies of *P. pinea*.

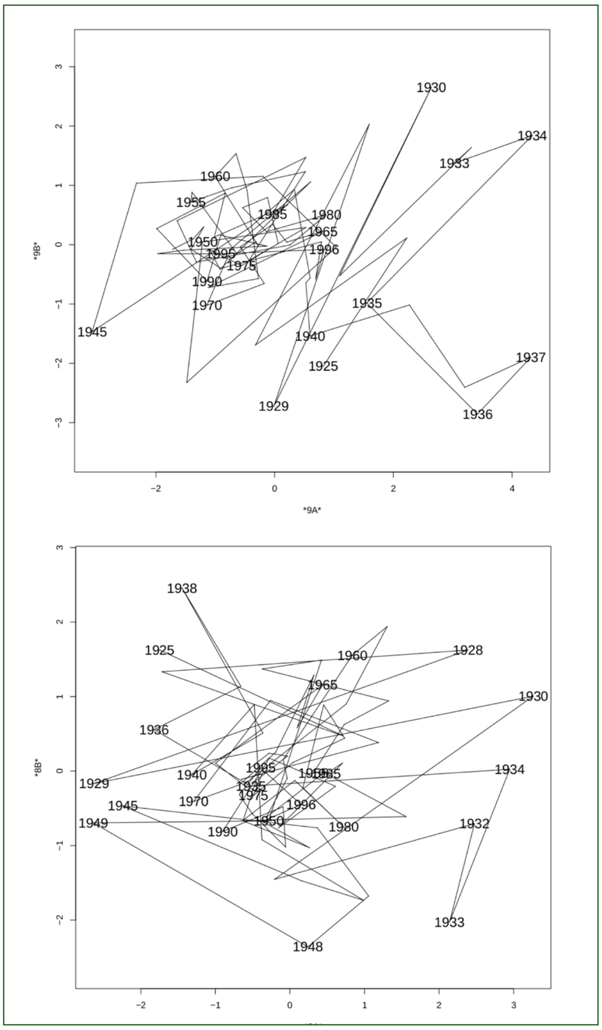


Table 4 - Correlation matrix of the four pairs of representative chronologies issued by HFC of the five chronologies of *P. pinea*. Here, the p-value for the correlations significance at 5% level is 0.22. The correlations significant at this level are in bold.

	9A	*9B*	*8A*	*8B*	*7A*	*7B*	*6A*	*6B*
9A	1.000							
9B	0.000	1.000						
8A	**0.368**	**0.930**	1.000					
8B*	0.158	-0.063	0.000	1.000				
7A	**0.981**	-0.196	0.179	0.168	1.000			
7B	0.010	0.051	0.051	0.149	0.000	1.000		
6A	**0.935**	-0.171	0.186	**0.206**	**0.951**	0.311	1.000	
6B	0.015	0.004	0.010	0.103	0.014	-0.043	0.000	1.000

Table 5 - Correlations of the five chronologies of *P. pinea* with both the four pair of representative time series issued by their HFC and the first three axes of their PCA. Here, the p-value for the correlations significance at 5% level is 0.22. The correlations significant at this level are in bold.

	9A	*9B*	*8A*	*8B*	*7A*	*7B*	*6A*	*6B*
San Rossore	**0.864**	-0.155	0.174	**0.230**	**0.877**	**0.267**	**0.917**	**0.399**
Cecina	**0.383**	**0.670**	**0.764**	**0.645**	0.244	**0.135**	0.274	0.074
Duna Feniglia	0.179	**0.751**	**0.764**	**-0.645**	0.028	-0.057	0.009	-0.059
Castelporziano	**0.851**	-0.158	0.166	0.147	**0.866**	**0.302**	**0.917**	**-0.399**
Circeo	0.727	-0.181	0.099	0.027	**0.748**	**-0.663**	0.505	0.039

vely of Duna Feniglia and Cecina in these years. In the lower graphics, corresponding to the node *8*, on the right side of the representative chronology (*8A*), the years appear in which Duna Feniglia and Cecina exhibit joint high values, like those from 1928 to 1934 (excluding 1929 and 1931), whereas on the left side 1929, 1945, and 1949 are found, with local minima. On the opposite, on the bottom side of the differences component *8B*, 1933, 1943, 1947, and 1948 turn out, with maxima for Duna Feniglia and not for Cecina, whereas on the top side 1938 and 1961 are found, with maxima for Cecina and minima for Duna Feniglia. Looking at the pattern of the years on the graphic, we may also say that the distribution is not particularly clustered around *8A* but rather uniformly distributed along *8B* too. This is in agreement with the low correlation between the two chronologies and suggests to keep them apart.

We may summarize these results by saying that the choice of three groups of chronologies seems reasonable, albeit some doubt may raise concerning Circeo, whose correlation with the representative chronology of group *6A* is medium (0.505). Indeed, the loss of correlation between Duna Feniglia and Cecina and between them and the representative chronology of group *7* is more than evident, thus preventing further aggregations around common chronologies. We may add that the representation of both the circle of correlations and of the years in the graphics associated to the last node *9* of the hierarchy (Fig. 6 left and 8 above), gives a reasonable oversight of the total structure of the data, albeit not optimal.

Discussion

The interpretation of the results given in the previous section was based on the rationale of PCA and HFC separately. Nevertheless, the comparison of the circles of correlations shown in Figures 2 and 6 proves that they are very similar, as well as the patterns of the years on the graphics of Figures 4 and 8 (above). Thus, the associated interpretation is nearly identical: the phenomenon may be described through at least three dimensions, corresponding to a group sharing a common chronology and two other independent ones. As for the years, their graphics are very similar too, by showing the extreme values far from the centroid nearly in the same directions.

Note that the very strong agreement between the first principal plane of PCA and the one of the last node of HFC results in general, albeit it is not theoretically proved. It may be argued that running only HFC might be sufficient to get the general information one usually examines in the first two dimensions of PCA, with the advantage to get the characters' hierarchy and the following nodes' factorial planes too.

As mentioned, common representations of the two methods' results may be realized, by projecting the principal components of the nodes issued by HFC as illustrative on the circles of correlations issued by PCA: they are represented in Figure 9, in practice the same of Figure 2 with the nodes' principal components projected on it as illustrative. This synthesizes well most of the aspects we dealt with in the results discussion. In general, the representative chronologies are situated as a weighted centroid within the chronologies gathered in the node, whereas the differences ones are orthogonal to their corresponding representative.

Looking at the circle on the left of Figure 9, we detect that *6A* and *7A* are oriented close to the first axis, to which *9A* is nearly coincident, and *8A* is oriented close to the second, with *9B* nearly coincident with it. This confirms the ability to HFC to mimic PCA, at least at its highest levels, i.e. the first two factors, despite of the constraints given by both the method itself (the iterated construction of principal time series PCAs) and the hierarchy. Actually, this is a sign of the quasi-optimality of HFC. Looking at the circle on the right of Figure 9, the nearly orthogona-

Figure 9 - Representation of the five chronologies of *P. pinea* (blue arrows) and the principal components of the nodes of the hierarchy HFC (red dotted arrows), on the circle of correlations on the plane spanned by the factorial axes 1-2 (left) and 2-3 (right) issued by their PCA.

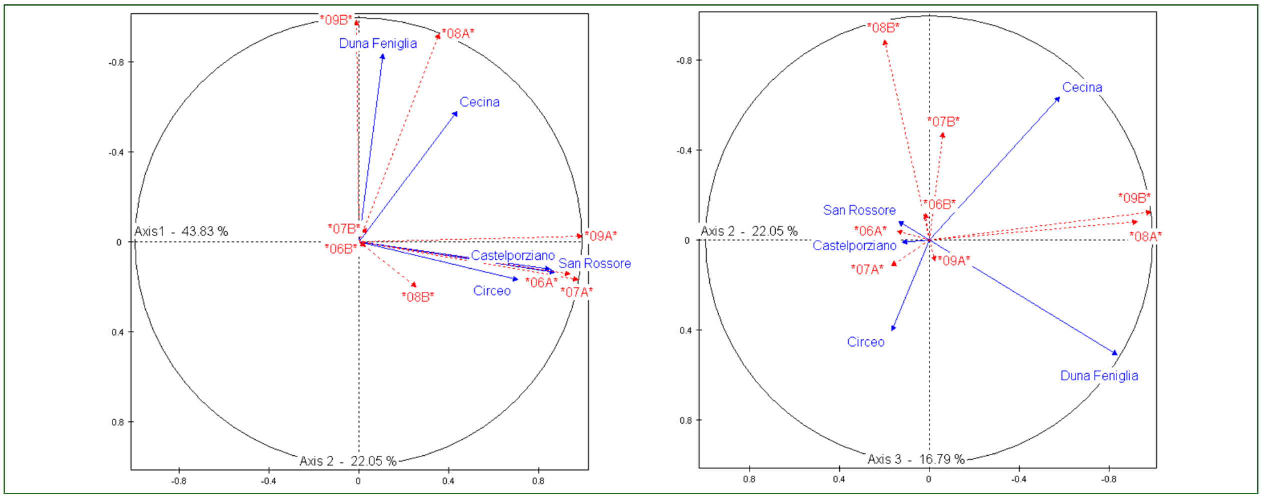


Table 6 - Correlations of the four pairs of representative time series issued by the HFC and the first three axes of PCA on the five chronologies of *P. pinea*. Here, the p-value for the correlations significance at 5% level is 0.22. The correlations significant at this level are in bold

	9A	*9B*	*8A*	*8B*	*7A*	*7B*	*6A*	*6B*
Axis 1	**0.995**	-0.011	**0.356**	**0.253**	**0.978**	0.033	**0.940**	0.029
Axis 2	0.025	**0.989**	**0.928**	-0.199	-0.170	0.063	-0.142	-0.017
Axis 3	-0.092	0.126	0.083	**0.892**	-0.115	**0.480**	0.040	0.118

lity between Duna Feniglia and Cecina was already commented: note also, the centroid position of *8A* in respect to them and the position of the differences component *8B*, oriented close to the third axis: the length of the latter may be a sign of the scarce evidence of a common signal. In order to better quantify the relations between representative chronologies and factors, we may look at Table 6. A very strong correlation of the first axis of PCA with the representative chronology of the last node *9A* (0.995) results. The correlation of the second axis with the differences principal component *9B* (0.989) and the correlation with the representative chronology *8A* are likewise very strong. Both *7A* and *6A* are very strongly correlated with the first axis; this is a sign of the stability of the chronology shared by the three sites encompassed in the same class, only little "disturbed" by the merging with the other two. Eventually, the two principal components of node *8* are strongly correlated with the second and third axes of PCA, respectively: along with the closeness of their inertias (1.169 and 0.831 respectively), this a clear sign of the independence of Duna Feniglia and Cecina from the group of three and between each other.

Both PCA and HFC are based on the correlation matrix and both may be used to get information which must lead to examine it according to the methods' aims. We did not show it until here, because the methods are usually run on large correlation matrices, whose reading would be quite difficult, with the aim to drive the user's attention directly on their most relevant aspects.

In Table 7 the correlation matrix between the chronologies is reported. Considering the number of years (79), the corresponding p-value for a correlation to be significant at 5% level is 0.22. Indeed, dealing with time-series, in which the auto-correlation (i.e. the correlation of a time-series with itself shifted by some time-lag) may not be negligible, this p-value must be considered with care, just as an indication of too lower correlations to be taken into account. Whereas zero- correlation would mean that no common signal shared by the series might be detected, larger correlation values may suggest the existence of a common signal.

Here, only four correlations are significant: those between San Rossore, Castelporziano, and Circeo, ranging within 0.681 and 0.448, and the weaker ones of Cecina with San Rossore and Castelporziano, barely above the threshold for significance. No significant correlation turns out for Duna Feniglia with the other chronologies. The current interpretation would lead to say that the first factor denotes a signal common to the three chronologies of the group, which may be very weakly shared by Cecina too, but not by Duna Feniglia.

It might be of interest to ascertain to what extent the representative chronologies may be consistently alternative to both PCC and the so-called standard chronologies (Peters et al. 1981), that is the average of all concerned chronologies, both on the statistical and the dendrochronological point of view. In Figure 10 the three of the group San Rossore, Castelporziano, and Circeo are compared: indeed, they are nearly identical, with a correlation between them of over 0.998, an outstanding performance, albeit the standard chronology – being an average – is the only one to get statistical properties.

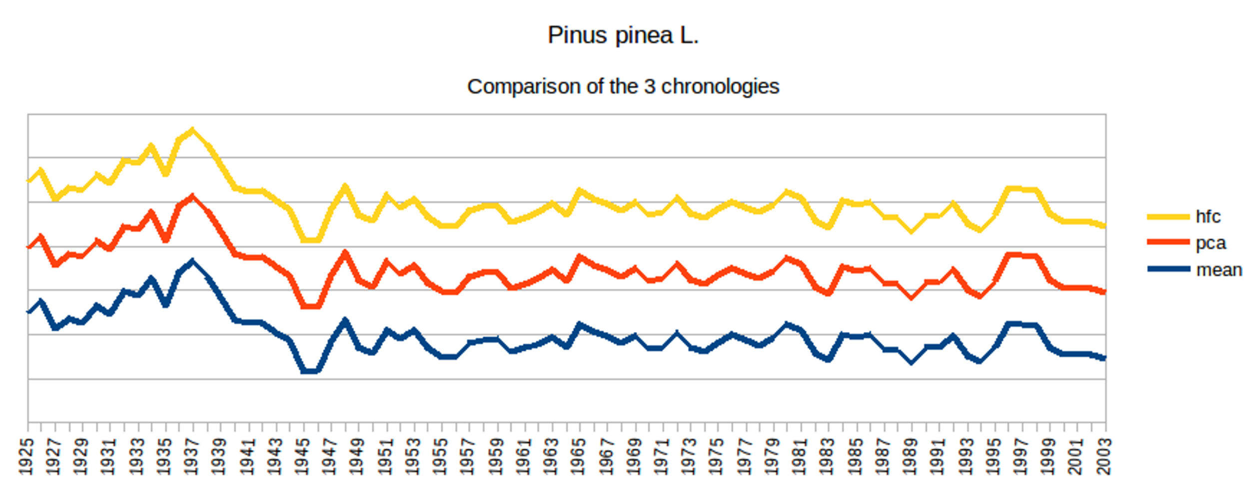
In order to provide an interpretation from the dendrochronology point of view, we may state that, within the time-span 1925-2003, three different chronologies come out, namely Cecina, Duna Feniglia, and one, shared by San Rossore, Castelporziano, and Circeo, which corresponds to the representative chronology *7A*. Moreover, the differences components *9B* and *8B* of the two highest nodes of the hierarchy result showing years in which the original chronologies are most different.

The three chronologies might be discussed considering both geographical proximity and environmental homogeneity of the study areas and their populations. Differences may be attributed to either intrinsic genetic diversity or very local environmental factors, triggering distinct responses by phenotypic plasticity, since no substantial differences in the macro-climatic envelopes of the coastal *P. pinea* stands are recorded in the study area, situated between 43°43' and 41°18' North. Considering the artificial origin of these populations, the almost complete lack of genetic variation in *P. pinea*, observed across the entire range of the species (Pinzauti et

Table 7 - Matrix of correlations between the 5 chronologies of *P. pinea*. Here, the p-value for the correlations significance at 5% level is 0.22. The correlations significant at this level are in bold.

	San Rossore	Cecina	Duna Feniglia	Castel-porziano	Circeo
San Rossore	1.000				
Cecina	**0.281**	1.000			
Duna Feniglia	-0.015	0.169	1.000		
Castelporziano	**0.681**	**0.222**	0.032	1.000	
Circeo	0.479	0.093	0.059	**0.448**	1.000

Figure 10 - Comparison between the representative chronology *7A* issued from HFC of the five chronologies of *P. pinea* by clustering San Rossore, Castelporziano, and Circeo (above), the first principal component issued from PCA run on these three chronologies only (center), and the standard chronology obtained by averaging the three chronologies (below).



al. 2012), leads to rule out any provenance-based response to the local environmental conditions due to genetic diversity, when heterogeneity of the chronologies is taken into account. However, since the species harbours a non-negligible amount of variation at adaptive traits (Vendramin et al. 2008), differences in soil conditions could be accounted for the different responses observed, due to the gradient of leaching, salinity, and to the heterogeneity in mineralogical composition of the sand dunes (from North to South: marly, potassic volcanic, clayey sands, see Carboni et al. 1994). In addition, the stands on Duna Feniglia are particularly affected by local specific environmental constrains. Enhanced topographic exposure to nearly constant sea winds, especially in the dry season, ongoing coastline retraction and salt-waterlogging on dune slacks, are likely to undermine the viability of the local populations and their growth pattern, in comparison to the stands growing in the other study sites. Moreover, populations of woody species of a whole array of contrasting plant communities from pure sclerophyllous Mediterranean evergreen aggregations to deciduous continental ones, may affect the local competition patterns of *P. pinea*.

The main issue of the application of HFC to *P. pinea* chronologies is to point out the pattern in common for the populations of San Rossore, Castelporziano, and Circeo, whereas the other two populations stand alone. Indeed, both Cecina and Duna Feniglia have been established on highly dynamic sites of littoral sand dunes, while the other stands, which share the same chronology, have been planted on planar areas further inland. In addition, Pinzauti et al. (2012) already quoted that Duna Feniglia pine forest is situated on a *tombolo*, a narrow sandy strip of dunes separating the sea from a lagoon, connecting the promontory of Monte Argentario (632 m.a.s.l.) with the Tuscanian coast, thus affected by salty water on both sides (Gabbrielli 1993) and exposed to both Northern and Southern quadrants dominating winds (Bellarosa et al. 1996). This particular situation sets Duna Feniglia really apart from the other sites. This is likely to account for the overall pattern of the proposed classification.

Conclusion

The highest correlation between the first two principal components of PCA of the total set of chronologies and the two corresponding to the last node issued by HFC, confirms the proximity of this method to PCA and its results. This proves that HFC is an effective alternative to PCA, with the additional ability of producing a hierarchy with consequent classifications and associated chronologies representative of the groups' chronologies and their differences. This is a real advantage for the dendrochronologist, since at each step of the hierarchy he/she may decide whether or not to consider it as a tentative chronology, carrying a signal common to those grouped together. Therefore, HFC seems particularly suited to address the complex patterns of relations between chronologies, at least for an exploratory study, better than other methods, requiring more relevant choices from the researcher: number of axes, rotations, etc. In fact, the definition of a common chronology for each obtained group is a significant advantage, not present in PCA.

As well as PCA, HFC may account for common signals, whose frequency may be highly variable, according to the data at hand. Thus, for the detection of a high-frequency common signal, further studies may be carried out with suitable data. Moreover, as PCA, HFC neither takes into account the variation of the correlation structure of the chronologies along time nor the ecological explanation of the results, which may depend upon environmental heterogeneity, hence from the joint analysis with other data sets. For their study, other more suitable exploratory multidimensional methods have already been used

(Fritts et al. 1970, Tardif et al. 2003) to compare chronologies and climatic factors; for the variation along time, some experiments have already been carried out (Camiz et al. 2010, Camiz and Roig 2011), but are still in progress.

We want to underline the exploratory nature of the analyses we dealt with here. By no means, to get more reliable results, confirmatory analyses, taking into account appropriate sampling and involving statistical tests, and further models are most suitable. Nevertheless, we are convinced that this way of studying may greatly help the researcher to organize his/her further steps in an optimal way to get better results.

References

Anderberg M.R. 1973 - *Cluster Analysis for Applications*. Academic Press New York.

Bellarosa R., Codipietro P., Piovesan G., Schirone B. 1996 - *Degradation, Rehabilitation and Sustainable Management of a Dunal Ecosystem in Central Italy*. Land Degradation & Development 7 (4): 297–311.

Benzécri J.-P. et coll. 1973 - *L'analyse des données*, 2. Dunod Paris.

Biondi F., Gershunov A., Cayan D.R. 2001 - *North Pacific Decadal Climate Variability since 1661*. Journal of Climate 14 (1): 5–10.

Boreux J., Naveau P., Guin O., Perreault L., Bernier J. 2009 - *Extracting a Common High Frequency Signal from Northern Quebec Black Spruce Tree-rings with a Bayesian Hierarchical Model*. Climate of the Past 5: 607–613.

Briffa K.R., Osborn T.J., Schweingruber F.H., Harris I.C., Jones P.D., Shiyatov S.G., Vaganov, E.A. 2001 - *Low-frequency Temperature Variations from a Northern Tree Ring Density Network*. Journal of Geophysical Research 106: 2929–2941.

Bry X. 1994 - *Analyses Factorielles Simples*. Economica Poche Paris.

Bunn A.G., Hughes M.K., Kirdyanov A.V., Losleben M., Shishov V.V., Berner L., Oltchev A., Vaganov E.A. 2013 - *Comparing Forest Measurements from Tree Rings and a Space-based Index of Vegetation Activity in Siberia*. Environmental Research Letters 8 (3): 035034.

Büntgen U., Frank D.C., Kaczka R.J., Verstege A., Zwijacz-Kozica T., Esper J. 2007 - *Growth Responses to Climate in a Multi-species Tree-ring Network in the Western Carpathian Tatra Mountains, Poland and Slovakia*. Tree Physiology 27: 689–702.

Camiz S., Denimal J.J. 2006 - *Classifications hiérarchiques factorielles de variables*. Revue des Nouvelles Technologies de l'Information E-6: 383–388.

Camiz S., Denimal J.J., Pillar V.D. 2006 - *Hierarchical Factor Classification of Variables in Ecology*. Community Ecology 7 (2): 165–179.

Camiz S., Maulucci R., Roig F. 2010 - *Exploratory Analysis Methods Applied to Dendrochronological Series*. In : Proceedings of Actas E-ICES 5, Malargüe (Mendoza, Argentina), 24-27 Octubre 2009, Papu, O. editor. Comisión Nacional de Energıa Atómica, Mendoza, Argentina: 61–67.

Camiz S., Pillar V.D. 2007 - *Comparison of Single and Complete Linkage Clustering with the Hierarchical Factor Classification of Variables*. Community Ecology 8 (1): 25–30.

Camiz S., Pillar V.D. 2018 - *Identifying the Informational/Signal Dimension in Principal Component Analysis*. Mathematics 6 (11): 269.

Camiz S. , Roig F.A. 2011 - *Evolutionary Analysis Applied to Tree-ring Chronologies Series*. In: Proceedings of Actas E-ICES 6, Malargüe (Mendoza, Argentina), 4-8 Octubre 2010. Comisión Nacional de Energí a Atómica, Buenos Aires: 30–37.

Carboni M., Palagi I., Palieri L., Raffi I., Sposato A. 1994 - *Dati preliminari sull'evoluzione geologica della fascia costiera del Lazio settentrionale durante il pliocene*. Memorie Descrittive della Carta Geologica d'Italia 49: 177–188.

Cook E.R., Kairiukstis A. 1992 - *Methods of Dendrochronology*. Kluwer Academic Publishers Dordrecht The Netherlands.

Denimal J.J. 2001 - *Hierarchical factorial analysis*. In: Proceedings of the 10th International Symposium on Applied Stochastic Models and Data Analysis. Université de Compiègne France.

Denimal J.J. 2007 - *Classification hiérarchique optimisée d'un tableau de mesures*. Revue de Statistique Appliquée 148 (2): 29–61.

Fowler A.M. 1988 - *Climatic Reconstructions from Tree Rings*. Weather and Climate 8: 33–45.

Frank D., Esper J. 2005 - *Characterization and Climate Response Patterns of a High-elevation, Multi-species Tree-ring Network in the European Alps*. Dendrochronologia 22 (2): 107–121.

Fritts H.C. 1976 - *Tree Rings and Climate*. Academic Press San Diego California.

Fritts H.C. 2012 - *Tree Rings and Climate*. Elsevier, Amsterdam.

Fritts H.C., Blasing T.J., Hayden, B.P., Kutzbach, J.E. 1970 - *Multivariate Techniques for Specifying Tree-Growth and Climate Relationships and for Reconstructing Anomalies in Paleoclimate*. Technical report, Laboratory of Tree-Ring Research, University of Arizona, Tucson.

Gabbrielli A. 1993 - *Origine delle pinete litoranee in Toscana*. In "Salvaguardia delle pinete litoranee" conference, 21-22 October 1993, Grosseto Italy: 13–20.

Gordon A.D. 1999 - *Classification*. Chapman and Hall/CRC Boca Ratón, Florida.

Gray S.T., Graumlich L.J., Betancourt J.L., Pederson G.T. 2004 - *A Tree-ring Based Reconstruction of the Atlantic Multidecadal Oscillation since 1567 A.D.*. Geophysical Research Letters 31 (12): 2–5.

Harman H.H. 1976 - *Modern Factor Analysis*. University of Chicago press Illinois.

Husson F., Lê S., Pagès J. 2017 - *Exploratory Multivariate Analysis by Example Using R*. Chapman and Hall/CRC Boca Ratón Florida.

Jolliffe I. 2002 - *Principal Component Analysis*. Springer Berlin.

Kendall M., Stuart A. 1973 - *The Advanced Theory of Statistics - volume 2: Inference and Relationship*. Griffin, London.

LaMarche V.C., Fritts H.C. 1971 - *Anomaly Patterns of Climate over the Western United States Derived from Principal Component Analysis of Tree-Ring Data*. Monthly Weather Review 99 (2): 138–142.

LaMarche V.C. 1974 - *Frequency-Dependent Relationships between Tree-Ring Series along an Ecological Gradient and some Dendroclimatic Implications*. Tree-Ring Bulletin 34: 1–20.

Lebart L., Piron M., Morineau A. 2006 - *Statistique exploratoire multidimensionnelle: Visualisations et inférences en fouille de données*. Dunod Paris.

Leland C., Pederson N., Hessl A., Nachin B., Davi N., D'Arrigo R., Jacoby, G. 2013 - *A Hydroclimatic Regionalization of Central Mongolia as Inferred from Tree Rings*. Dendrochronologia 31 (3): 205–215.

Lerman I.C. 1981 - *Classification et analyse ordinale des données*. Dunod Paris.

Litton C.D., Zainodin H.J. 1991 - *Statistical Models of Dendrochronology.* Journal of Archaeological Science 18 (4): 429–440.

MacQueen J. 1967 - *Some Methods for Classification and Analysis of Multivariate Observations*. In: Proceedings of the fifth Berkeley Symposium on Mathematical Statistics and Probability. Oakland, California, vol.1: 281–297.

Mazza G., Cutini A., Manetti M.C. 2014 - *Site-specific Growth Responses to Climate Drivers of* Pinus Pinea *L. Tree Rings in Italian Coastal Stands*. Annals of Forest Science 71 (8): 927–936.

Nelson B.D. 2001 - *Variable Reduction for Modeling Using Proc Varclus*. In: Proceedings of the Twenty-Sixth Annual SAS Users Group International Conference, SAS Institute, Cary North-Carolina.

Papadopoulos A. 2016 - *Tree-ring Patterns and Climate Response of Mediterranean Fir Populations in Central Greece*. Dendrochronologia 40: 17–25.

Papadopoulos A., Tolika K., Pantera A., Maheras P. 2009 - *Investigation of the Annual Variability of the Aleppo Pine Tree-Rings Width: The Relationships with the Climatic Conditions in the Attica Basin*. Global Nest Journal 11 (4): 583–592.

Patskoski J., Sankarasubramanian A., Wang H. 2015 - *Reconstructed Streamflow Using SST and Tree-ring Chronologies over the Southeastern United States*. Journal of Hydrology 527: 761–775.

Peters K., Jacoby G.C., Cook E.R. 1981 - *Principal Components Analysis of Tree-ring Sites*. Tree-Ring Bulletin 41: 1–19.

Pinzauti F., Sebastiani S., Budde K., Fady B., González-Martí nez S., Vendramin G. 2012 - *Nuclear Microsatellites for* Pinus Pinea *(Pinaceae), a Genetically Depauperated Tree, and their Transferability to* Pinus Halepensis. American Journal of Botany: e362–e365.

Piovesan G., Biondi F., Bernabei M., Di Filippo A., Schirone B. 2005 - *Spatial and Altitudinal Bioclimatic Zones of the Italian Peninsula Identified from a Beech (*Fagus Sylvatica *L.) Tree-ring Network*. Acta Oecologica 27 (3): 197–210.

Piraino S., Camiz S., Di Filippo A., Piovesan G., Spada F. 2013 - *A Dendrochronological Analysis of* Pinus Pinea *L. on the Italian Mid-tyrrhenian Coast*. Geochronometria 40 (1): 77–89.

Rohli R. V., Vega A. J. 2018 - *Climatology*. Jones & Bartlett Learning Burlington Massachusetts.

Saint George S., Meko D., Evans M. 2008 - *Regional Tree Growth and Inferred Summer Climate in the Winnipeg River Basin, Canada, since AD 1783*. Quaternary Research 70: 158–172.

Speer J.H. 2010 - *Fundamentals of Tree-ring Research*. University of Arizona Press, Tucson.

Stafasani M., Toromani E. 2015 - *Growth-Climate Response of Young Turkey Oak (*Quercus Cerris *L.) Coppice Forest Stands along Longitudinal Gradient in Albania.* South-east European Forestry 6 (1): 25–38.

Tardif J., Camarero J.J., Ribas M., Gutiérrez E. 2003 - *Spatiotemporal Variability in Tree Growth in the Central Pyrenées: Climatic and Site Influences*. Ecological Monographs 73 (2): 241–257.

Touchan R., Shishov V.V., Tychkov I.I., Sivrikaya F., Attieh J., Ketmen M., Stephan J., Mitsopoulos I., Christou A., Meko D.M. 2016 - *Elevation-layered Dendroclimatic Signal in Eastern Mediterranean Tree Rings*. Environmental Research Letters 11 (4): 044020.

Tukey J. 1977 - *Exploratory Data Analysis*. Addison-Wesley Reading Massachusetts.

Vendramin G.G., Fady B., González-Martınez S.C., Hu F.S., Scotti I., Sebastiani F., Soto Á., Petit R. J. 2008 - *Genetically Depauperate but Widespread: the Case of an Emblematic Mediterranean Pine*. Evolution 62 (3): 680–688.

Vigneau E., Qannari E., Sahmer K., Ladiray D. 2006 - *Classification de variables autour de composantes latentes*. Revue de Statistique Appliquée 54 (1): 27–45.

Whittaker R.E. 1973 - *Handbook of Vegetation Science. V Ordination and Classification of Vegetation*. Junk The Hagu.