Bulletin of the *Transilvania* University of Braşov
Series II: Forestry • Wood Industry • Agricultural Food Engineering • Vol. 5 (54) No. 1 - 2012

CLUSTER ANALYSIS IN PUBESCENT OAK TAXA FROM SERIES LANUGINOSAE: A CASE STUDY

C.M. ENESCU[1] N. ŞOFLETEA[1] A.L. CURTU[1]

Abstract: *The main purpose of this study was to compare the different clustering techniques in order to identify the one with the best discriminating power among oak species. By using two groups of trees, one corresponding to pedunculate oak and the other one to pubescent oak, Ward's Method with Manhattan distances provided the best separation between the two oak species. By contrast, different results were obtained for the two taxa from series Lanuginosae (pubescent oak and Italian oak), due mainly to their similarities in leaf and fruit traits. In conclusion, by using a wide variety of clustering combinations within STATISTICA software, no separation between pubescent oak and Italian oak was achieved.*

Key words: *Cluster analysis, Lanuginosae, pubescent oak, Italian oak.*

1. Introduction

In general, the term cluster analysis refers to a broad series of techniques used for classification of individuals [8]. In particular, cluster analysis is a common statistic tool used in studies aimed to discriminate between closely related oak species (genus *Quercus, Fagaceae*) [1], [4], [12], [14], [15], or to group different populations of the same species according to their geographic origin [2], [7], [13]. Its goal is to sort different objects (i.e. trees) into homogenous groups. Therefore, on the basis of similarity, the trees within a group (cluster) will be related to one another and unrelated to the trees from other groups.

Nowadays, there are several clustering techniques provided by different statistical software packages and sometimes selecting the best one is complicated.

In the case of the oaks (*Quercus* spp.), many amalgamation (linkage) rules and distances were used for cluster analysis. The most frequent linkage rules were: *i*) Complete Linkage [3], [10]; *ii*) Single Linkage or the Nearest Neighbour [11]; *iii*) Unweighted Pair Group Method with Arithmetic Mean [6], [24] and *iv*) Ward's Method [19], [23], while the usual distances used were: *i*) the Euclidean [6], [7], [10], [15]; *ii*) the Square Euclidean [17]; *iii*) the Manhattan [1] and *iv*) the Mahalanobis [13] ones.

Given that the oaks constitute one of the most species-rich genera and the fact that the identification of species is often complicated by the overlapping morphological traits, the taxonomical status of some oak species is sometimes a subject of debate. If the taxonomical status of pubescent or downy oak (*Quercus*

[1] Dept. of Silviculture, *Transilvania* University of Braşov.

pubescens Willd.) According to different taxonomical classifications, Italian oak (Q. virgiliana Ten.) is sometimes considered as a distinct species [5], [20], sometimes it is not [18].

2. Objectives

Firstly, the objective of this study was to compare different clustering techniques in the case of two easily identified by morphological criteria oak species, namely pedunculate oak (*Quercus robur* L.) and pubescent oak (*Q. pubescens*). Secondly, the best clustering technique was used in the case of the two taxa from series *Lanuginosae*, in which morphological differentiation is difficult or questionable.

3. Materials and Methods

The methodology for evaluating the leaf morphology was also used in other similar studies [9], [16], [22]. So, five leaves per tree were chosen and 14 leaf descriptors were assessed: five dimensional characters - *lamina length* (LL), *petiole length* (PL), *lobe width* (LW), *sinus width* (SW), *length of lamina at largest width* (WP), two counted variables - *number of lobes* (NL) and *number of intercalary veins* (NV), two observed variables - *abaxial laminar pubescence* (PU), evaluated according to the grading system from 1 (no pubescence) to 6 (dense hairness) with an microscope (x30) and *basal shape of the lamina* (BS), scored as an index varying from 1 to 9 and five transformed variables - *obversity* (OB), *petiole ratio* (PR), *lobe depth ratio* (LDR), *percentage venation* (PV) and *lobe width ratio* (LWR). In addition, for the two taxa from series *Lanuginosae* the *length of fruit peduncle* (Lp) was also assessed.

In order to achieve the main objective of this study, the first data set for analysis was composed of fifty pedunculate oaks and fifty pubescent oaks. The reason for choosing these two oak species is the obvious difference regarding the leaf morphology. Particularly, compared to pubescent oak, pedunculate oak has a longer lamina, but a shorter petiole, an ear-like basal shape of the lamina and no hairs on the abaxial part of the leaf [21]. So, the proper cluster technique should be able to provide a good separation of the two oaks. Secondly, the best clustering techniques resulted from the first data set were used to separate the two taxa from series *Lanuginosae* (75 pubescent oak trees and 25 Italian oaks). The mean values of all descriptors, except the calculated ones are given in Table 1.

Mean values for descriptors Table 1

Variable	1st set		2nd set	
	Q. robur	*Q. pubescens*	*Q. pubescens*	*Q. virgiliana*
PU	1	4	4	4
BS	8	4	4	4
NL	11	10	10	9
NV	4	2	2	2
LL [mm]	114	81	82	98
PL [mm]	5	14	14	15
LW [mm]	34	29	29	31
SW [mm]	14	12	12	15
WP [mm]	71	44	44	54
Lp [mm]	x	x	5	25

Cluster analyses were carried out with STATISTICA software v.8.0. Several clustering amalgamation rules (Complete Linkage, Single Linkage, Unweighted pair-group average, Weighted pair-group average and Ward's Method) and all available distances within the software were applied to the data. In total, 35 combinations (five linkage rules multiplied with seven distances) were made.

4. Results and Discussions

As revealed by the tree diagram (Figure 1), Ward's Method with Manhattan (City-Block) distances provided the best separation between *Quercus pubescens* and *Q. robur*. It can be seen that the cluster analysis delimited two distinct groups which correspond to the two oak species.

Similarly, two homogenous groups were also observed in the case of other related oak species, namely *Q. robur* and *Q. petraea* [15], or *Q. virginiana* and *Q. geminata* [4]. Moreover, good results were also obtained for the following combinations: Ward's Method with Euclidean distances, Weighted pair-group Average with Manhattan distances or Unweighted pair-group average with percent disagreement.

On the contrary, bad results were obtained when Single Linkage was used as amalgamation rule. In this case, the cluster analysis produced more than two groups and no species assignment could be observed (Figure 2).

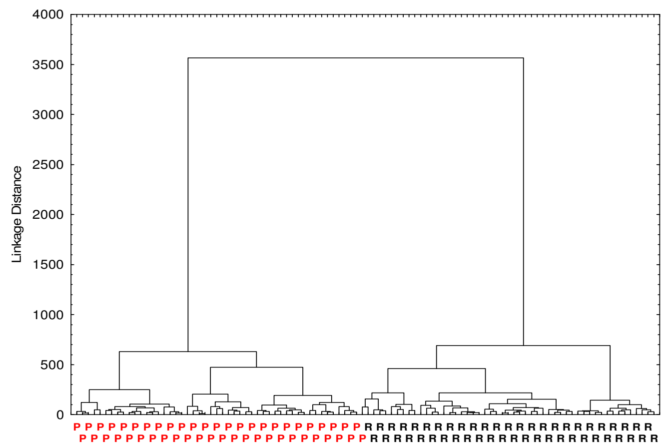

Fig. 1. *Dendrogram of the first data set* (*Ward's Method*).
Abbreviations: P - Q. pubescens, R - Q. robur

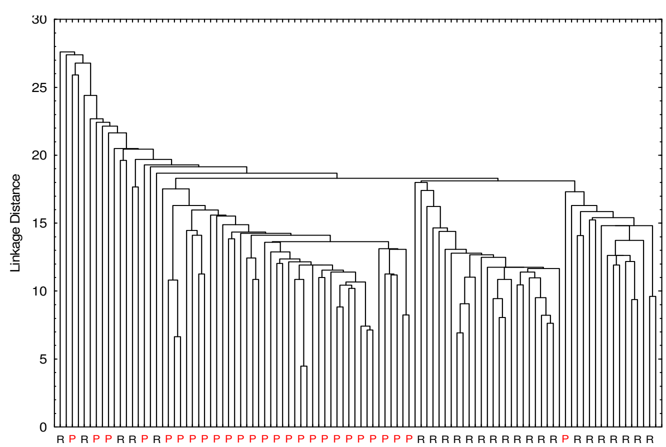

Fig. 2. *Dendrogram of the first data set* (*Single Linkage*).
Abbreviations: P - Q. pubescens, R - Q. robur

Based on these results, the best combination of cluster analysis was applied to the second set of data. Interestingly, for this data set, two mixed groups were observed (Figure 3).

These findings support the hypothesis that the two taxa from series *Lanuginosae* cannot be recognized as different species [18], due mainly to their similarity in leaf and fruit traits [21]. This is in agreement with the results from Principal Component Analysis conducted for 931 pubescent oaks from Romania, where only one morphological group was observed [10].

Furthermore, regarding the same one hundred analysed trees from series *Lanuginosae* no significant differences were observed in allele frequencies at seven microsatellite loci (unpublished data).

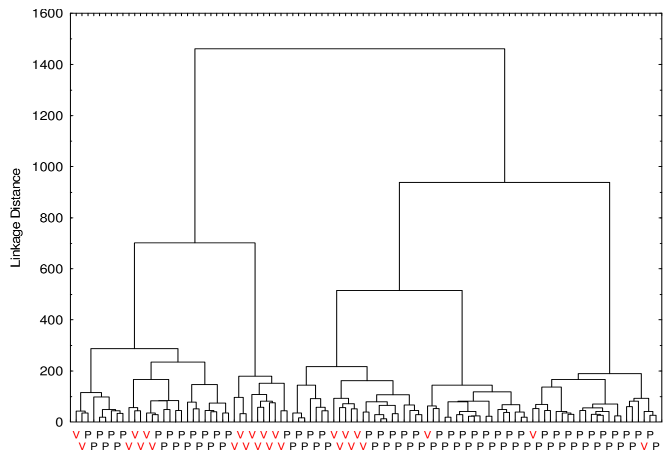

Fig. 3. *Cluster diagram of the second data set.*
Abbreviations: P - Q. pubescens, V - Q. virgiliana

5. Conclusions

By using all the clustering techniques provided by STATISTICA software no separation between the two taxa from series *Lanuginosae* was possible. This was mainly due to their similarities in leaf and fruit traits, the only descriptor which somehow separates them being the *length of the peduncle* (Lp) [10].

On the other hand, Ward's Method with Manhattan distances proved to be the best choice for discriminating between pedunculate oak and pubescent oak.

Our findings from this statistical-morphological survey correlated with those from genetic investigations done for the same trees by using seven nuclear microsatellite loci (unpublished data) suggest that *Q. virgiliana* is rather an intra-specific unit of *Q. pubescens* than a separate species.

Acknowledgements

This work was supported by CNCS-UEFISCDI, project number PN II-RU-TE-73/2010. Cristian Mihai Enescu was supported by the Sectoral Operational Programme Human Resources Development (SOP HRD), financed from the European Social Fund and by the Romanian Government under the contract number POSDRU/88/1.5/S/59321.

References

1. Aas, G.: *Taxonomical Impact of Morphological Variation in Quercus*

robur and Q. petraea: A Contribution to the Hybrid Controversy. In: Annals of Forest Sciences **50** (1993) No. 1, p. 107s-113s.

2. Broshtilov, K.: *Quercus robur L. Leaf Variability in Bulgaria*. In: PGR Newsletter **147** (2006), p. 64-71.
3. Bruschi, P., Vendramin, G.G., et al.: *Morphological and Molecular Diversity among Italian Populations of Quercus patraea* (*Fagaceae*). In: Annals of Botany **91** (2003) No. 6, p. 707-716.
4. Cavender-Bares, J., Pahlich, A.: *Molecular, Morphological, and Ecological Niche Differentiation of Sympatric Sister Oak Species, Quercus virginiana and Q. geminata* (*Fagaceae*). In: American Journal of Botany **96** (2009) No. 9, p. 1690-1701.
5. Ciocârlan, V.: *Flora ilustrată a României* (*Illustrated Flora of Romania*). Bucureşti. Ceres Publishing House, 2000.
6. Cristofolini, G., Crema, S.: *A Morphometric Study of the Quercus crenata Species Complex* (*Fagaceae*). In: Botanica Helvetica **115** (2005), p. 155-167.
7. Cuza, P.: *Variabilitatea frunzelor stejarului pufos* (*Quercus pubescens Wild.*) *în Republica Moldova* (*Leaf Variability of pubescent oak (Quercus pubescens Wild.*) *in Republic of Moldova*). In: Mediul Ambient **5** (2010) No. 53, p. 7-14.
8. Dytham, C.: *Choosing and Using Statistics: A Biologist's Guide*. Second Edition, Cornwall, TJ International Ltd Publishing House, 2003.
9. Enescu, C.M., Chesnoiu, E.N., et al.: *Leaf Morphology in Quercus robur L. Genetic Resources across Romania*. In: Bulletin of the *Transilvania* University of Braşov, Series II, Vol. 3 (52), 2010, p. 47-54.
10. Enescu, C.M., Şofletea, N., Curtu, A.L.: *Leaf Multivariate Analysis in Four Romanian Oak Species.* In: Proceedings of the 4th International Conference of Young Scientists, Prague, Czech Republic, Czech University of Life Sciences, April 28, 2011, p. 57-66.
11. Ferreira-Dias, S., Valente, D.G., Abreu, J.M.F.: *Pattern Recognition of Acorns from Different Quercus Species Based On Oil Content and Fatty Acid Profile*. In: Grasas y Aceites **54** (2003) No. 4, p. 384-391.
12. Finkeldey, R., Mátyás, G.: *Genetic Variation of Oaks* (*Quercus spp.*) *in Switzerland. 3. Lack of Impact of Postglacial Recolonization History on Nuclear Gene Loci.* In: Theoretical and Applied Genetics **106** (2003), p. 346-352.
13. Franjić, J., Liber, Z., Škvorc, Ž.: *Morphological and Molecular Differentiation of the Croatian Populations of Quercus pubescens Willd.* (*Fagaceae*). In: Acta Societatis Botanicorum Poloniae **75** (2006) No. 2, p. 123-130.
14. Jensen, J., Larsen, A., et al.: *Hybridization between Quercus robur and Q. petraea in a Mixed Oak Stand in Denmark*. In: Annals of Forest Sciences **66** (2009) No. 7, p. 706-717.
15. Kelleher, C.T., Kelly, D.L., Hodkinson, T.R.: *Species Status, Hybridisation and Geographic Distribution of Irish Populations of Quercus petraea (Matt.) Libel. and Q. robur L.* In: Watsonia **25** (2004), p. 83-97.
16. Kremer, A., Dupouey, J.L., et al.: *Leaf Morphological Differentiation between Quercus robur and Quercus petraea is Stable across Western European Mixed Oak Stands.* In: Annales des Sciences Forestières **59** (2002), p. 777-787.
17. Minotta, G., Degioanni, D.: *Naturally Regenerated English oak (Quercus robur L.) Stands on Abandoned*

Agricultural Lands in Rilate Valley (*Piedmont Region, NW Italy*). In: iForest - Biogeosciences and Forestry **4** (2011), p. 31-37.

18. Nixon, K.C.: *Infrageneric Classification of Quercus* (*Fagaceae*) *and Typification of Sectional Names.* In: Annales des Sciences Forestières **50** (1993) No. 1, p. 25-34.
19. Rushton, B.S.: *Quercus robur L. and Quercus petraea* (*Matt*.) *Liebl.: A Multivariate Approach to the Hybrid Problem, 1. Data Acquisition, Analysis and Interpretation*. In: Watsonia **12** (1978), p. 81-101.
20. Schwarz, O.: *Quercus L.* In: *Flora Europaea,* Tutin, T.G., et al. (Eds). Cambridge. Cambridge University Press, 1993, p. 72-76.
21. Şofletea, N., Curtu, A.L.: *Dendrologie* (*Dendrology*). Braşov. *Transilvania* University Publishing House, 2007.
22. Şofletea, N., Enescu, C.M., Curtu, A.L.: *Small-Scale Morphological Descriptors Analusis in Four Romanian Oak Stands Reported to Series Lanuginosae Simk*. In: Bulletin of the *Transilvania* University of Braşov, Series II, Vol. 4 (53) No. 2, 2011, p. 77-84.
23. Ufnalski, K.: *Teleconnection of 23 Modern Chronologies of Quercus robur and Q. petraea from Poland*. In: Dendrobiology **55** (2006), p. 51-56.
24. Zhang, X., Korpelainen, H., Chunyang, L.: *Microsatellite Variation of Quercus aquifolioides Populations at Varying Altitudes in the Wolong Natural Reserve of China*. In: Silva Fennica **40** (2006) No. 3, p. 407-415.