Biometrical Letters
Vol. 48 (2011), No. 1, 67-81

Discriminant analysis for Kraft's classes of trees

Jolanta Grala-Michalak[1], Katarzyna Kaźmierczak[2]

[1]Faculty of Mathematics and Computer Science, Adam Mickiewicz University, Umultowska 87, 61-614 Poznań, Poland, e-mail: grala@amu.edu.pl
[2]Department of Forest Management, Poznan University of Life Sciences, Wojska Polskiego 71c, 60-625 Poznań, Poland, e-mail: kasiakdendro@wp.pl

SUMMARY

The paper presents results of discriminant analysis for Kraft's classes of trees. Kraft's classification is based on a tree's position in the stand's social structure and its crown development and extent. Belonging to a given social class reflects the position of a tree in a stand, and through this, its growth potential. The aim of the analysis was to select variables which mostly determined the Kraft class of a tree and to construct discriminant functions which assign data well to Kraft's classes.

Key words: discriminant analysis, LDA, PCA, Scots pine

1. Introduction

There are many cases in which foresters classify trees. In the 19*th* century, Kraft (1884) (in Assmann, 1961) created what would become one of the most widely used tree classifications. This classification is based on a tree's position in the stand's social structure and its crown development and extent. Kraft recognized the following classes of trees:

- predominant trees with exceptionally well-developed crowns
- dominant trees, forming the main stand as a rule with relatively well-developed crowns
- low co-dominant trees; crown shape is still normal and hence the trees are similar to those in the second tree class in this respect, yet they are relatively weakly developed and restricted often already with the onset of degeneration

The classes 1–3 are called the dominant stand.

- dominated trees, with crowns more or less dying back, restricted on all sides or on two sides, or with one-sided development
    - intermediate trees, essentially free of canopy cover with restricted lateral crown growth
    - partially overtopped crowns, the upper crown free, the lower crown under canopy cover
- entirely overtopped trees
    - with crowns capable of growth
    - with dead crowns

The classes 4–5 are called the suppressed stand. Belonging to a given social class reflects the position of a tree in a stand, and through this, its growth potential.

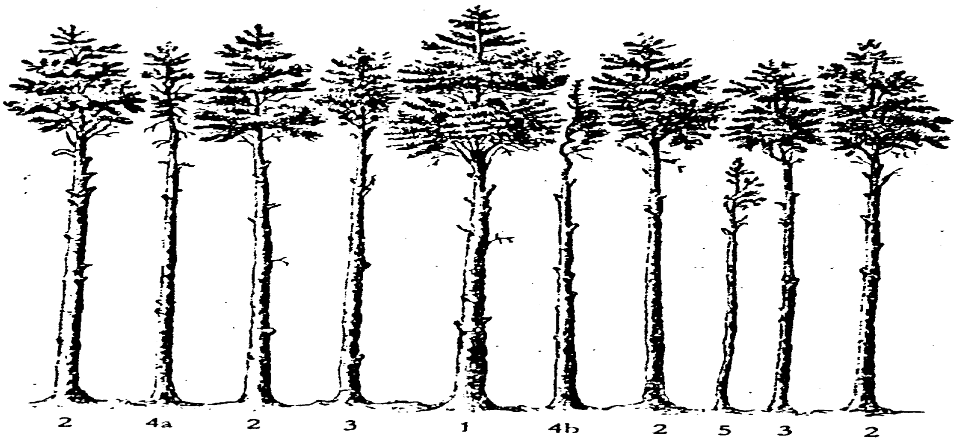

Figure 1. Kraft's classes (1884 in Assmann 1961)

The aim of the analysis is to choose the variables which mostly determined the Kraft's class of a tree (1, 2, 3, 4a, 4b, 5a) and to construct discriminant functions which assign data well to Kraft's classes. Knowledge about the position of a tree in the forest may help draw conclusions about the forest management in the past and apply these to the future.

In the past discriminant analysis has been applied to analyse measurements on trees, but in another context. For example, Sahunalu *et al.* (1994) carried out research on the relationship between soil and plants in Thailand's dry dipterocarp forest. Wyant *et al.* (1986) studied the dependence between the percentage of fire damage to crowns and the mortality of trees in a Colorado pine and fir forest. Reitberger *et al.* (2008) used a discrimination method for tree recognition on 3D full waveform LIDAR data from the Bavarian Forest National Park. Niche differences in four species of Galium were quantified by using discriminant function analysis of site characteristics including biotic variables (Mann and Shugart 1984). Lewis and Rice (1990) estimated the risk of erosion on forest lands. Discriminant analysis was used to classify rain forest types in Costa Rica (Thessler *et al.*, 2008). Discriminant analysis, probit analysis and logit analysis were compared for the prediction of individual overstory tree mortality in northern hardwood stands in Wisconsin (Monserud, 1976). Blackard and Dean (1999) applied discriminant analysis in predicting forest cover types from cartographic variables. Here this statistical methodology is used to Kraft's classification of the trees.

2. Experimental material

The experimental material included selected results for 200 pine-tree trunks derived from 8 stands. All stands from which the experimental test trees were derived grew on the fresh mixed coniferous forest sites situated in the Zielonka Experimental Forest District. Sample trees followed the methodology developed by Draudt. The same calendar growth period from 1989 to 1993 was adopted for each tree. The purpose of this assumption was to rule out the effect of additional factors such as site, climate and meteorology on the increment. Prior to felling their social class was established according to the criteria proposed by Kraft. In the presented study the following traits were measured:

- age of sample trees (w),
- tree height (h) measured in m

- 5-year increment in height (Zh_5) in m
- double bark thickness (K) – measured in cm at a height of 1.3 m from the base of the tree
- breast height diameter inside bark ($d_{1.3.}$) – its diameter measured in cm at the height of 1.3 m from the base of the tree
- 5-year increment in breast height diameter (Zd_5) in cm
- tree basal area ($g_{1.3}$) in m^2 – the area of a circle with a diameter equal to the breast height diameter
- 5-year basal area increment (Zg_5) in m^2
- tree volume (V), i.e. section-based volume in m^3; the stem is divided into sections of identical length (in this case 1 m) – the volume of each complete section is established as the volume of a cylinder with diameter measured at mid-length of a given section, while the volume of the last incomplete section is calculated using a formula for the volume of a cone; these calculated volumes of individual sections are summed
- 5-year volume increment (Zv_5)
- breast height form factor ($f_{1.3}$) equal to $V/(g_{1.3}*h)$
- volume growth intensity coefficient (i_5) equal to $Zv_5/ g_{1.3}$
- tree slenderness (s) defined as the ratio of height in m to breast height diameter in cm (h/ $d_{1,3}$)

Most of the analysed traits were determined for standing trees, although Zh_5, V, Zv_5, $f_{1.3}$, i_5 were measured on felled trees.

3. Statistical methods

The problem of classification arises when an investigator makes a number of measurements on an individual and wishes to classify the individual into one of a finite number of categories, but cannot do so directly from the measurements. For example when a tree has been felled and transported to a sawmill it is possible to measure some of the above variables but there is no information about the position of that tree in the stand social structure. We are interested in finding some linear functions of several number of variables representing the features measured on a individual tree which are the best predictors of the assignment to groups.

The method of discrimination analysis is widely described in Krzyśko (1990, 2000), Koronacki and Ćwik (2005) and Anderson (2003) and in many other books. The kind of analysis used in this paper is described in the new book by Krzyśko at al. (2008). Information on the practical use of STATISTICA is can be easily found in Stanisz (2007).

We take into consideration 13 variables: *w*, *h*, Zh_5, *K*, $d_{1,3}$, Zg_5, *V*, Zv_5, $f_{1.3}$, i_5, *s*. A grouping variable 'Kraft' was adopted to discriminate the groups 1, 2, 3, 4a, 4b and 5a. The calculations were performed using STATISTICA.

We let $\mathbf{Y}_i = (\mathbf{X}_{i1}, \mathbf{X}_{i2}, ..., \mathbf{X}_{in_i})'$ be the random sample from the *i*-th (i = 1, ..., *K*) p-variate normally distributed population (Kraft's group) from the set $\{\pi_1, \ldots, \pi_K\}$ and $n_1 + n_2 + \ldots + n_K = n$. The maximum likelihood estimators of the mean value of predictors $\boldsymbol{\mu}_i$ and the variance-covariance matrix $\boldsymbol{\Sigma}_i$ of π_i has the form (respectively)

$$\overline{\mathbf{X}}_i = \frac{1}{n_i}\sum \mathbf{X}_{ij} \quad \text{and} \quad \mathbf{S}_i = \frac{1}{n_i - 1}\sum (\mathbf{X}_{ij} - \overline{\mathbf{X}}_{ij})(\mathbf{X}_{ij} - \overline{\mathbf{X}}_{ij})' .$$

From the whole training sample $\mathbf{X} = (\mathbf{Y}_1', \mathbf{Y}_2', ..., \mathbf{X}_n')'$ we calculate $\overline{\mathbf{X}} = n^{-1}\sum n_i \overline{\mathbf{X}}_i$, the variance between sample's matrix $\mathbf{B} = \sum n_i (\overline{\mathbf{X}}_i - \overline{\mathbf{X}})(\overline{\mathbf{X}}_i - \overline{\mathbf{X}})'$ and the variance inter sample's matrix $\mathbf{W} = \sum (n_i - 1)\mathbf{S}_i$.

We are interested in finding the set of constant vectors $\{\boldsymbol{a}_i : i = 1, \ldots, K\}$ which maximalizes the expression $(n - K)\mathbf{a}'\mathbf{B}\mathbf{a}/(K-1)\mathbf{a}'\mathbf{W}\mathbf{a}$ on condition that $(n-K)^{-1}\mathbf{a}_i'\mathbf{W}\mathbf{a}_j = \delta_{ij}$ where $\boldsymbol{a} = (\boldsymbol{a}_1', \ldots, \boldsymbol{a}_K')'$ is a matrix built on the $\boldsymbol{a}_i$'s and δ_{ij} is a Kronecker's delta (i.e. 1 if $i=j$ and 0 if $i \neq j$). This criterion means that new variables (discriminant functions) $u_i = \boldsymbol{a}_i'\boldsymbol{x}$ are uncorrelated, with variances equal to 1 for every *i*.

The first discriminant function u_1 is related to the first (i.e. the largest) eigenvalue λ_1 of the matrix $(n-K)(K-1)^{-1}\mathbf{W}^{-1}\mathbf{B}$, the second is related to λ_2, and so forth. The vectors $\boldsymbol{a}_i$ are the eigenvectors related to the λ_i's, where $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_s > 0$ are the solutions of the determinant equation $(\mathbf{B} - (K-1)(n-K)^{-1}\lambda_i \mathbf{W})\mathbf{a}_i = \mathbf{0}$, for i = 1,…,s.

In order to make scales of prior variables equal, the $\boldsymbol{u}_k$'s should be standardized in this manner $u_k = a_{k1}(x_{ij1} - \bar{x}_{i1}) + \ldots + a_{kp}(x_{ijp} - \bar{x}_{ip})$ for k = 1,…,s. The absolute values of the coefficients $\boldsymbol{a}_{kr}$'s divided by the r-*th* root of the diagonal element of $(n-K)^{-1}\mathbf{W}$ show the contribution of prior variables $\boldsymbol{x}_1,\ldots,\boldsymbol{x}_p$ to discrimination of groups by the k-th discriminant variable $\boldsymbol{u}_k$.

In the classification process a discriminant variable u_k is not useful if λ_k is not significantly different from zero (Krzyśko at al. 2008). First we test the hypothesis that all eigenvalues are equal to 0, next that they all are apart from the first, etc. This procedure is continued until the first time the hypothesis is not rejected. Then we state that from this value the remaining (n-s) eigenvalues are equal to 0. As the test statistic we apply Wilks' lambda of the form

$$\Lambda_d = \prod_{i=d+1}^{s} \frac{1}{1+\lambda_i} \quad \text{where } d = 0, 1, \ldots, p\text{-}1,$$

which has asymptotic χ^2 distribution (more in Krzyśko, 1990, Koronacki and Ćwik, 2005, Krzyśko at al., 2008). Its values are shown in Table 1. To evaluate the prior variable we use the partial Wilks' lambda associated with the unique contribution of the respective variable to the discriminatory power of the model.

When we use the linear classifier (in LDA) then we assign observations to the class k having the smallest value of the function

$$\hat{d}_k(x) = \sum_{i=1}^{s} [a_i'(x - \bar{x}_k)]^2 \quad \text{for } k = 1, 2, \ldots, K$$

where $\bar{\mathbf{x}}_k$ is the vector of means for the k-th class. (Krzyśko *et al.*, 2008).

4. First results

At first, we include all variables and all cases in the model (see Table 1). Wilks' statistic called lambda is based on the clustering of points in p-dimensional space around the centroid. Its value is in [0,1] and 0 denotes the best power of discrimination. The partial lambda measures the influence of the variable on discrimination and is equal to the lambda statistic for all variables over the lambda statistic for all but the one variable.

Table 1. Choice of variables for discriminant functions (good candidates in bold)

Var.	Wilks' Lambda	Partial Lambda	Standardized coefficients for u_1	Standardized coefficients for u_2	p-level	R^2
W	**0.092**	**0.893**	**-1,07755**	-1,17180	**0.00087**	**0.938**
h	**0.094**	**0.874**	**0,71741**	**-2,08813**	**0.00016**	**0.989**
Zh_5	0.085	0.964	**0,39139**	-0,12656	0.24656	0.844
K	0.085	0.970	0,04208	0,12650	0.35264	0.810
$d_{1\cdot3}$	**0.091**	**0.898**	**0,37484**	**4,17564**	**0.00143**	**0.998**
Zd_5	**0.097**	**0.849**	-0,08578	**3,31982**	**0.00001**	**0.970**
$g_{1.3}$	0.090	0.912	**1,53303**	2,04069	**0.00475**	**0.998**
Zg_5	**0.092**	**0.892**	-0,09533	**-4,07365**	**0.00084**	**0.986**
V	**0.091**	**0.902**	-0,09533	**-3,02553**	**0.00205**	**0.994**
Zv_5	**0.096**	**0.853**	**-0,77727**	**3,20533**	**0.00002**	**0.972**
$f_{1.3}$	0.090	0.915	0,06057	0,53291	**0.00590**	0.714
i_5	**0.108**	**0.761**	**0,99521**	**-2,65707**	**0.00000**	**0.945**
s	**0.101**	**0.815**	**-0,77326**	1,97743	**0.00000**	**0.939**

According to Krzyśko i in. (2008) we should build our model on the basis of all variables significant at the level 0.01 level (w, h, $d_{1\cdot3}$, Zd_5, $g_{1.3}$, Zg_5, V, Zv_5, $f_{1.3}$, i_5, s) which contributes most (have largest absolute value of coefficients) in the first discriminant function (w, h, Zh_5, $d_{1\cdot3}$, $g_{1.3}$, Zv_5, i_5, s) or in the second one (h, $d_{1\cdot3}$, Zd_5, Zg_5, V, Zv_5, i_5). Obviously, it is not possible to take all three variables i_5, Zv_5, $g_{1.3}$, or all three of s_5, h, $d_{1,3}$, or all four of $f_{1.3}$, V, $g_{1.3}$, h in the same model because of the definitions of the quantities.

The quantity R^2 measures how much the variance of the grouping variable is explained by the variables in the model. Although the adding a new variable to the model makes R^2 bigger, one should check if the new variable is redundant, i.e. duplicates the influence of another one in the model.

5. The choice of variables

We are interested in making the shortest list of variables included in the model but with a great discriminatory power. Unfortunately, there are a lot of models satisfying the above conditions. In order to find the best of them the PCA method will be applied (see Jolliffe, 1982). We find the lines called factor axes generated by a set of orthogonal eigenvectors of the covariance matrix **W** (Table 2). Only the three largest eigenvalues are related (jointly) to the great part of the total variation, so the projections of 13 variables on two-factor planes (1x2) and (2x3) will show the importance and relations between them.

Table 2. The largest eigenvalues of the data covariance matrix and the percentage of total variance explained

Eigenvalue	7.51	3.12	1.06	0.54	0.25	0.19
%	57.76	24.02	8.18	4.15	1.95	1.46

Figure 2 presents 13 variables as the end points of eigenvectors projected on to the two-factor plane. With the exception of $f_{1.3}$ and K the remaining variables are highly correlated with the factor axes, because they are located almost on the unit circle. The uncorrelated pairs of variables are presented as the orthogonal pairs of vectors. The points Zh_5, s, w and Zv_5 (or Zg_5) lie on the vertices of a square (see Figure 2). Variables located on the same diameter of the circle are negatively correlated with the value –1 and both should not be entered in the same model. We get two models: (s, w) and (Zh_5, Zg_5, w) which are seemed to be adequate to the data. Additionally $f_{1.3}$ is ortogonal to the remain variables on (2x3)-plane in Figure 3, so $f_{1.3}$ and i_5, (lying orthogonally to $f_{13.}$) ought to be included in our model. The projection on the (1x3)-plane looks similarly to Figure 2, so is not presented here.

Let us consider two models: (s, w, $f_{1.3}$, i_5) and (Zh_5, Zg_5, w, , $f_{1.3}$, i_5). The eigenvalues of the first model are equal (approximately) to 3.162, 0.067, 0.026

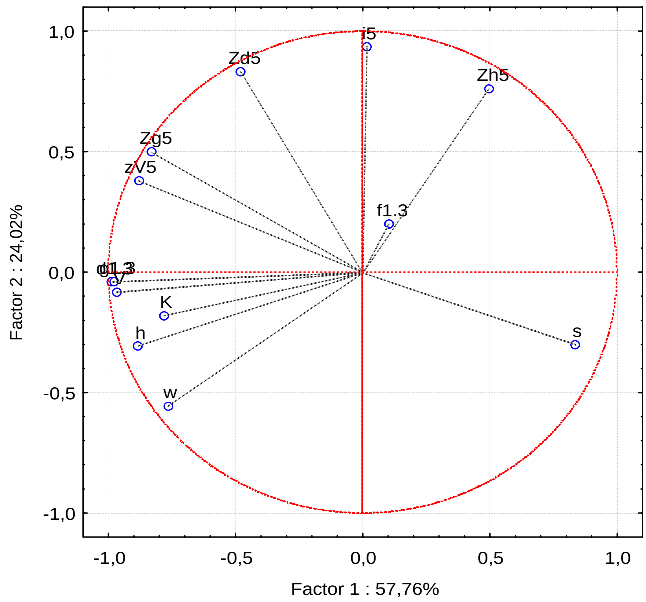

Figure. 2. Projection of the variables on the two-factor plane (1x2)

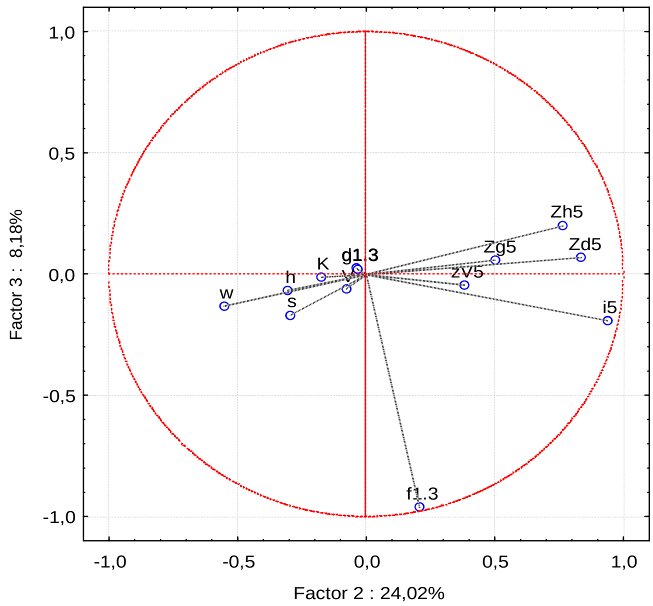

Figure 3. Projection of the variables on the two-factor plane (2x3)

and lower, so it is appropriate to take one discriminant function because very small eigenvalues have a negligible influence on the grouping variable. To go into detail the χ^2 test rejected the second and subsequent functions as not significant at the level 0.01 (Table 3). The first discriminant function is good enough because of Wilks' lambda value. The second model has two discriminant functions significant at the level 0.01 (see Table 4).

Table 3. Characteristics of discriminant functions in the model ($s,w, f_{1.3},i_5$)

Discriminant function	Eigenvalue	Wilks' lambda	χ^2	Degrees of freedom	p-level
1	3.162*	0.218	295.31	20	0.000
2	0.067	0.908	18.65	12	0.097
3	0.026	0.970	6.00	6	0.615

*-significant at 0.01

Table 4. Characteristics of discriminant functions in model (Zh5, Zg5, w, , f1.3 , i5 **)**

Discriminant function	Eigenvalue	Wilks' lambda	χ^2	Degrees of freedom	p-level
1	1.876*	0.250	268.55	25	0.000
2	0.347*	0.718	64.16	16	0.000
3	0.028	0.967	6.45	9	0.694

*-significant at 0.01

In the first model the only one discriminant function

$$u_1 = 0.170196 \cdot w + 0.732688 \cdot i_5 + 0.056777 \cdot f_{1.3} + 1.063779 \cdot s$$

explains 97% of total variation (see the last row in Table 5). The variable s has the greatest influence on discriminant function u_1 because of the obtained absolute value of standardized coefficients.

In the second model discriminant functions

$$u_1 = 0.442695 \cdot w + 0.274062 \cdot i_5 + 0.022642 \cdot f_{1.3} + 0.758828 \cdot Zh_5 + 0.865548 \cdot Zg_5$$

$$u_2 = 2.19449 \cdot w + 1.16317 \cdot i_5 - 0.37051 \cdot f_{1.3} + 0.90217 \cdot Zh_5 - 1.04235 \cdot Zg_5$$

explain 83% and 15% of the total variation, respectively. The variables Zg_5 has the greatest influence on discriminant function u_1, and the variable w on u_2.

Table 5. Standardized coefficients of canonical variables

Variable	$(s,w, f_{1.3},i_5)$ U_1	$(Zh_5, Zg_5, w, f_{1.3}, i_5)$ u_1	u_2
W	0.170196	0.442695	2.19449
I_5	0.732688	0.274062	1.16317
$f_{1.3}$	0.056777	0.022642	-0.37051
S	1.063779	-	-
Zh_5	-	0.758828	0.90217
Zg_5	-	0.865548	-1.04235
Eigenvalue	3.162347	1.875704	0.34747
Cum. Prop.	0.969751	0.831097	0.98506

Obviously there are observations far from the centroid of its Kraft's group, and so the classification cannot be done easily. This is implied by the fact that some trees can change their class during their life, if for example the dominating trees close to them die are dead or are felled. Because of this (as one can see in Figure 4) the transition from observations in a certain class to the next class is continuous along a straight line parallel to the u_1 axis. This means that there is no important difference between the type of classifiers; linear, quadratic or elliptical discriminant functions will give the similar result. Therefore only linear discriminant functions are considered.

In Figure 4 groups are placed along a straight line, because the second discriminant function is negligible. In the second model, the groups in Figure 5 lie along a parabolic line. This means that the first discriminant function orders the means for groups exactly according to their names whereas the second distinguish the extreme groups from those in the centre, as one can see from the coefficients in Table 6.

Table 6. Means of canonical variables for Kraft's groups

Kraft's classes	u_1 in $(s, w, f_{1.3},i_5)$	u_1 in $(Zh_5, Zg_5, w, f_{1.3}, i_5)$	u_2 in $(Zh_5, Zg_5, w, f_{1.3}, i_5)$
1	-2.53452	2.46925	-0.63631
2	-1.09142	0.65534	-0.01444
3	0.20682	-0.25359	0.67153
4a	1.66872	-1.41986	0.11239
4b	2.78263	-1.94038	-0.34888
5a	4.60742	-2.75888	-2.00517

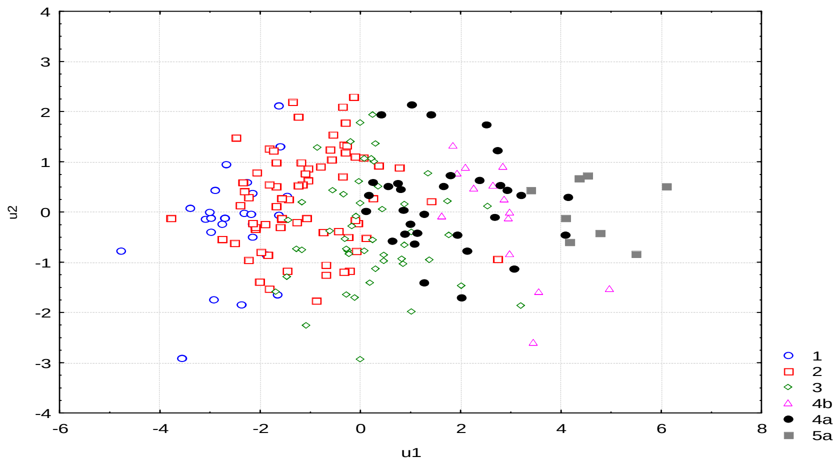

Figure 4. Discrimination by u_1 and u_2 in the model (s, w, $f_{1.3}$, i_5)

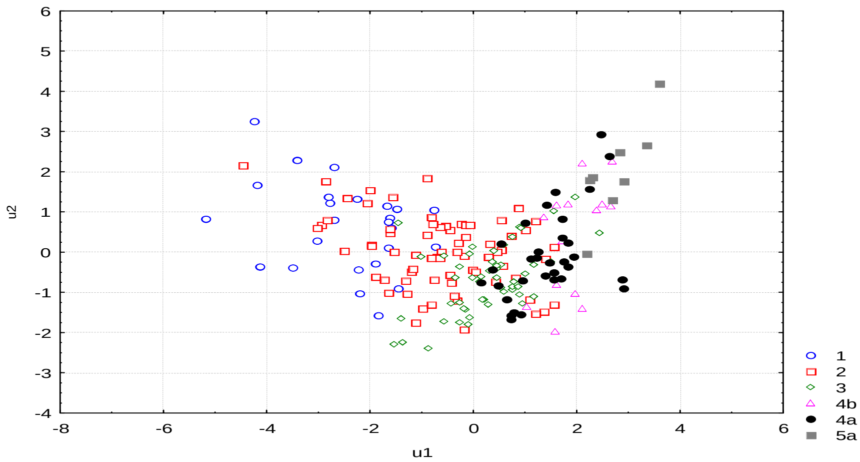

Figure 5. Discrimination by u_1 and u_2 in model (Zh_5, Zg_5, w, $f_{1.3}$, i_5)

6. Results of classification

As we do not have a test sample (additional data) independent from the learning sample (the basis for constructing the discriminant functions) the cross-validation procedure will be performed to verify which of the two models is better. Let us divide the sample into two parts of 100 cases. 50% of every Kraft's class data will create the learning sample to estimate parameters of a model and the remains will be classified according to that model. This division was simulated 50 times at random. A priori classification probabilities are proportional to group sizes and equal to 0.125, 0.35, 0.255, 0.16, 0.07 and 0.04, respectively. The effects of simulations are presented in Tables 7 and 8.

Table 7. Average percentage of correct classification in the model (s, w, $f_{1.3}$, i_5)

Kraft's class	Classified as 1	Classified as 2	Classified as 3	Classified as 4a	Classified as 4b	Classified as 5a
1	**48.33**	19.14	0.00	0.00	0.00	0.00
2	18.33	**75.43**	20.60	4.29	5.94	0.00
3	0.42	26.43	**52.00**	7.14	16.88	0.00
4a	0.00	1.86	26.00	**20.71**	41.88	14.00
4b	0.00	0.00	0.20	27.86	**28.75**	8.00
5a	0.00	0.00	0.00	16.43	0.31	**56.00**

Table 8. Average percentage of correct classification in the model (Zh5, Zg5, w, f1.3 , i5)

Kraft's class	Classified as 1	Classified as 2	Classified as 3	Classified as 4a	Classified as 4b	Classified as 5a
1	**55.42**	15.71	1.40	0.00	0.00	0.00
2	35.42	**60.29**	27.80	5.00	14.69	0.00
3	0.00	19.14	**60.0**	12.14	17.50	3.00
4a	0.00	2.00	22.80	**15.71**	45.94	23.00
4b	0.00	0.00	4.60	25.00	**19.38**	20.00
5a	0.00	0.00	0.00	10.00	4.38	**52.00**

The model (s, w, $f_{1.3}$, i_5) is slightly better than the model (Zh_5, Zg_5, w, $f_{1.3}$, i_5). Neither is good enough to distinguish classes 4a and 4b classes, because there are significantly less observations in 4b than in 4a. Although the size of the last (5a) is 10 (the least), this group is identifiable with the use to identify with the help of both models.

7. Conclusions

The percentage of dominating trees in the stand gives us the important information about the productivity of that stand. Obviously the social position of a tree in the stand may naturally change over time, more frequently in

younger stands than in older ones. The discriminant functions given in this paper allow us to provide Kraft's classification continuously in time and choose the best moment for removing trees. Although it happens that we assing a tree to a group in the neighbourhood of the proper one, it is relatively rare that we misclassify in a group, lying far from the true group.
The considered models are equivalent in the sense that s measures the ratio of height to breast height diameter at a certain moment in time, whereas Zh_5 and Zg_5 measure 5-year increment in height and 5-year basal area increment.

Knowing the four values (of w, $f_{1.3}$, i_5 and s) for a tree we are able to assign it to Kraft's classes properly with a probability at least 50% if we treat classes 4a and 4b as one we get 60.8% in the first model and 69.9% in the second one. It can be done by calculating the value of the function u_1 in model (s, w, $f_{1.3}$, i_5). The results of classification using the model (Zh_5, Zg_5, w, $f_{1.3}$, i_5) are slightly worse in the sense that we estimate one parameter more and classify cases using two discriminant functions, but the results are similar.

Though grouped according to just a very simple criterion $\hat{d}_k(x)$ trees within the same Kraft's class prove to be quite homogenous (as far as many dendrometrical features are concerned), while there was significant variation between the classes. Each Kraft's class groups trees with similar growth potential.

REFERENCES

Anderson T.W. (2003): An Introduction to Multivariate Statistical Analysis. Wiley, New York.

Assmann E. (1968).:The Forest Productivity. (in Polish), PWRiL, Warsaw.

Blackard J.A., Dean D.J. (1999): Comparative accuracies of artificial neural networks and Discriminant Analysis in Predicting Forest Cover Types from Cartographic Variables. Computers and Electronics in Agriculture 24: 131-151.

Draper N.R., Smith H. (1966): Applied Regression Analysis. Wiley, New York.

Jolliffe J.T., (1982): A note on the use of Principal Components in Regression. Journal of the Royal Statistical Society, Series C, 31(3): 300-303.

Koronacki J., Ćwik J. (2005): Statistical Learning Systems. (in Polish), WNT, Warsaw.

Krzyśko M. (1990): Discriminant Analysis. (in Polish), WNT, Warsaw.

Krzyśko M. (2000): Multivariate Statistical Analysis. (in Polish), UAM - Poznań.

Krzyśko M., Wołyński W., Górecki T., Skorzybut M. (2008): Learning Systems. Patters Recognition, Cluster Analysis, Reduction of Dimensionality. (in Polish), WNT, Warsaw.

Lewis J., Rice R.M. (1990): Estimating Erosion Risk on Forest Lands Using Improved Methods of Discriminant Analysis. Water Resour. Res. 26(8): 1721-1733.

Mann L.K., Shugart H.H. (1983): Discriminant Analysis Some East Tennessee Forest Herb Niches. Plant Ecology 52(2): 77-89.

Monserud R.A. (1976): Simulation of Forestn Tree Mortality. Forest Science 22(4): 438-444.

Reitberger J., Schnörr Cl., Krzystek P. and Stilla U. (2008): 3D Segmentation of Full Waveform LIDAR Data for Single Tree Detection Using Normalized Cut www.isprs.org/proceedings/ XXXVII/congress/3_pdf/12.pdf

Sahunalu P., Dhanmanonda P., Khemnark Ch. (1994): Discriminant Analysis of Soil and Plant Relationships in Dry Dipterocarp Forest. Thai Journal of Forestry, Bangkok.

Stanisz A. (2007): The Easy Accesible Statistical Course with STATISTICA PL and Medicine Examples, Vol. 3. Multidimensional Analyses, (in Polish), Statsoft, Kraków.

Thessler S., Sesnie S., Ramos Bendafia Z.S., Ruokolainen K., Tomppo E., Finegan B. (2008): Using k-nn and Discriminant Analyses to Classify Rain Forest Types in a Landsat TM Image over Northern Costa Rica. Remote sensing of environment 112(5): 2485-2494.

Wyant J. G., Omi P. N. and Laven R. D. (1986): Fire Induced Tree Mortality in a Colorado Ponderosa Pine/Douglas-fir Stand, Forest Sci. 32(1): 49-59.