Point distribution form model for spruce stems (*Picea abies* [L.] Karst.)

M. KŘEPELA

Czech University of Agriculture, Faculty of Forestry, Prague, Czech Republic

ABSTRACT: The paper deals with the construction of a point distribution form model for spruce stems. This model is based on the principal components analysis of variance-covariance matrix formed for the Procrustes residuals. The calculation of full Procrustes co-ordinates, the principal components, is demonstrated on an example of a spruce experimental plot at premature age, and a point distribution model is constructed for the first three components. The parameters of the model are evaluated in relation to Konšel's (Kraft's) tree classes, normality of their classification is tested, maxima and minima are demonstrated on actual trees. The complete stem shape analysis of all four samples is also provided. A special model is constructed for these samples and the course of the parameters of this model is graphically represented.

Keywords: Norway spruce (*Picea abies* [L.] Karst.); stem shape; Procrustes analysis; principal components analysis; point distribution form model

The examination of stem form has a long tradition in dendrometry and is of great importance for the construction of volume tables, assortment tables, tables of simple growth curves as well as growth models. In the last 20 years so called geometric methods of shape description have been developed in connection with computer tomography and computer image processing. These methods are used for wood testing just as Metriguard Inc. company's products show (www. metriguard.com). Their advantage consists in a clear definition of concepts and possibility to compare shapes using multidimensional statistical methods. The shape of configuration $\boldsymbol{x}_i$ is all the geometric information about $\boldsymbol{x}_i$ that is invariant under location, rotation and isotropic scaling (Euclidean similarity transformations see Fig. 1 – DRYDEN, MARDIA 1998).

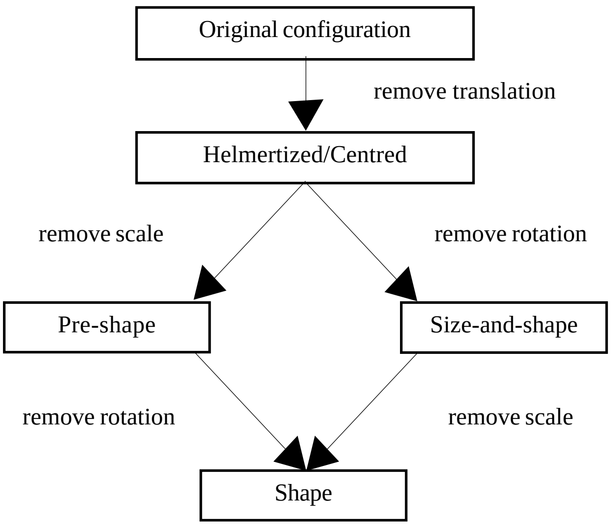

Fig. 1. Scheme of shape evaluation (by GOODALL, MARDIA 1992 in DRYDEN, MARDIA 1998), adapted

COOTES et al. (1992, 1994) used *the principal components analysis* (PCA) to develop a *point distribution form model* (PDM) where the principal components model for shape and Procrustes residuals are used.

Use of PCA is an application of a morphometric pioneer recommendation: "We must learn from the mathematician to eliminate and to discard; to keep the type in mind and leave the single case, with all its accidents, alone" (THOMPSON 1917).

METHODOLOGY

ALIGNING THE TRAINING SET

Given n independent configurations $\boldsymbol{x}_1,...., \boldsymbol{x}_n$, $\boldsymbol{x}_i = (x_{i1},..., x_{ik})^T$ in $\boldsymbol{C}$. The configuration is a set of landmarks on a particular object. Configurations are centred $\boldsymbol{x}_i^* \boldsymbol{1} = 0$, $\boldsymbol{x}_i^*$ denotes the transpose of the complex conjugate of $\boldsymbol{x}_i^*$ $\boldsymbol{1}$ is vector of ones). The figures are registered

This study was supported by Grant Project No. 526/01/0922 funded by the Grant Agency of the Czech Republic.

to $\boldsymbol{x}^P_1,...., \boldsymbol{x}^P_n$, by complete generalized *Procrustes analysis*.

KENT (1994) in DRYDEN and MARDIA (1998) defines the *full Procrustes mean shape* $\hat{\boldsymbol{\mu}}$ as the eigenvector corresponding to the highest eigenvalue of the complex sum of squares and products matrix

$$\boldsymbol{S} = \sum_{i=1}^{n} \boldsymbol{x}_i \boldsymbol{x}_i^* / (x_i^* x_i) = \sum_{i=1}^{n} \boldsymbol{z}_i \boldsymbol{z}_i^*$$

where $\boldsymbol{z}_i = \boldsymbol{x}_i / \|\boldsymbol{x}_i\|$, $i = 1, \ldots, n$, are the pre-shapes.

The full Procrustes fits or *full Procrustes co-ordinates* of $\boldsymbol{x}_i, \ldots, \boldsymbol{x}_n$ are

$$\boldsymbol{x}_i^P = \boldsymbol{x}_i^* \hat{\boldsymbol{\mu}} \boldsymbol{x}_i / (\boldsymbol{x}_i^* \boldsymbol{x}_i), \qquad i = 1, \ldots, n$$

where each $\boldsymbol{x}_i^P$ is the full Procrustes fit of $\boldsymbol{x}_i$ onto $\hat{\boldsymbol{\mu}}$. Calculation of the full Procrustes mean shape can also be obtained by taking the arithmetic mean of the full Procrustes co-ordinates, i.e.

$$\frac{1}{n} \sum_{i=1}^{n} \boldsymbol{x}_i^P$$

has the same shape as the full Procrustes mean shape $\hat{\boldsymbol{\mu}}$.

Procrustes co-ordinates were calculated by means of "tpsRegr" v. 1.20 by ROHLF (1998).

Complex numbers are used for simplifying the calculation. There are several programmes available on the address: http://life.bio.sunysb.edu/morph/software.html. The other presented calculations can be processed by the common statistical software, factually the SPSS Advanced Statistics 7.5 was used.

POINT DISTRIBUTION FORM MODEL

(According to COOTES et al. [1992, 1994]).

Let us take $\hat{\boldsymbol{\mu}}$ and $\boldsymbol{x}_i^P$ as the real vectors

$$\boldsymbol{x}_i^P = (x_{i1}, y_{i1}, \ldots, x_{ik}, y_{ik})^T$$

For each $\boldsymbol{x}_i^P$ we will calculate the deviation $d\boldsymbol{x}_i^P$ from the full Procrustes mean $\hat{\boldsymbol{\mu}}$, so

$$d\boldsymbol{x}_i^P = \boldsymbol{x}_i^P - \hat{\boldsymbol{\mu}}$$

We can then calculate the variance-covariance matrix, $\boldsymbol{S}$, using

$$\boldsymbol{S} = 1/n \sum_{i=1}^{n} d\boldsymbol{x}_i^P \, (d\boldsymbol{x}_i^P)^T$$

The modes of variation of the points of the shape are described by the unit eigenvectors of $\boldsymbol{S}$, $\boldsymbol{p}_i$ ($i = 1$ to $2k$) so that

$$\boldsymbol{S}\boldsymbol{p}_i = \lambda_i \boldsymbol{p}_i$$

(where λ_i is the i'th eigenvalue of $\boldsymbol{S}$, $\lambda_i \geq \lambda_{i+1}$)

$$\boldsymbol{p}_i^T \boldsymbol{p}_i = \mathbf{1}$$

It can be shown that the eigenvectors of the variance-covariance matrix corresponding to the highest eigenvalues describe the most significant modes of variation in the variables used to derive the variance-covariance matrix, and that the proportion of total variance explained by each eigenvector is equal to the corresponding eigenvalue. Most of the variation can usually be explained by a small number, t, of modes. One method for calculation t would be to choose the smallest number of modes so that the sum of variance explained would be a sufficiently large proportion of λ_T, the total variance of all the variables, where

$$\lambda_T = \sum_{i=1}^{2k} \lambda_i$$

The i'th eigenvector affects point k in the model by moving it along a vector parallel to $(dx_{i,}, dy_{i,})$, which is obtained from the k'th pair of elements in $\boldsymbol{p}_i$.

Any shape in the training set can be approximated using the mean shape and weighted sum of these deviations obtained from the first t modes

$$\boldsymbol{x}_i^P = \hat{\boldsymbol{\mu}} + \boldsymbol{P}\boldsymbol{b}_i$$

where $\boldsymbol{P} = (\boldsymbol{p}_1, \boldsymbol{p}_2, \ldots \boldsymbol{p}_t)$ is the matrix of the first t eigenvectors,

$$\boldsymbol{b}_i = (b_1, b_2, \ldots, b_t)$$

is a vector of weights for each eigenvector, the eigenvectors are orthogonal, $\boldsymbol{P}^T \boldsymbol{P} = \boldsymbol{I}$ so

$$\boldsymbol{b}_i = \boldsymbol{P}^T (\boldsymbol{x}_i^P - \hat{\boldsymbol{\mu}}) \qquad (1)$$

PRACTICAL EXAMPLES

COMPLETE EXPERIMENTAL PLOT

The stems from the experimental plot Doubravčice 1 are taken as an example for using the point distribution model. This plot is situated in the School Forest District of the Czech University of Agriculture at Kostelec nad Černými lesy. In 1966 the plot was felled, and dendrometric characteristics of individual trees were investigated. The plot lies in the forest type 2K3 (acid beech-oak stand), its area is 0.5 ha, mean age 70 years, upper stand height $h_{100} = 26.3$ m, mean stand height $h_g = 23.3$ m, mean stand diameter $d_g = 21.6$ cm. The pure spruce stand with density $\rho = 1.0$ was of concern (KŘEPELA et al. 2001).

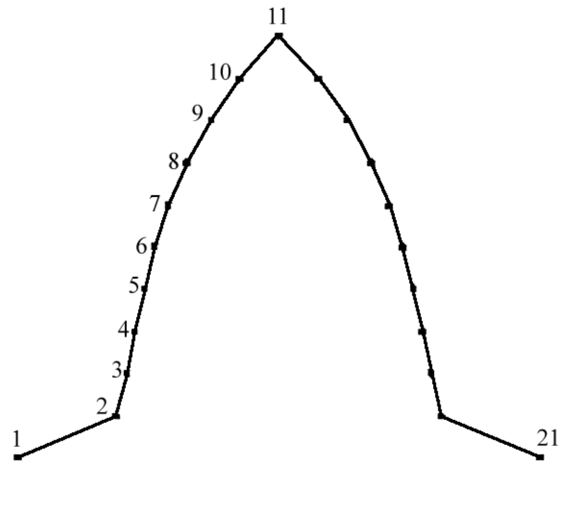

Fig. 2. 21-point model of spruce stem

From this plot 481 stems were used for investigations. The stem lengths and sections were measured with measuring tape to the nearest 1cm, as well as the diameters with metal rule to the nearest 0.5 mm in 2m sections and at the height of 1.3 m above ground (KORF 1972). The initial section has the height and diameter of the stump. These data were recounted for sections the length of which was increasing by 1/10 of tree height. All the diameters were measured over bark. The marginal points of such chosen sections form 21 landmarks on the morphological curve and their x- and y- co-ordinates create configuration $\boldsymbol{x}_i$ (see Fig. 2).

From this configuration the full Procrustes mean shape $\hat{\boldsymbol{\mu}}$ and full Procrustes co-ordinates $\boldsymbol{x}_i^P$ were calculated by means of the programme TpsRegr Version 1.20 (ROHLF 1998). The variance-covariance matrix $\boldsymbol{S}$ was also calculated as well as its eigenvalues and eigenvectors were found out.

Survey of the first three eigenvalues that include 99% variability is presented in Table 1.

Table 1. Eigenvalues of the variance-covariance matrix from the set of spruce stems and proportional expression of variability explained by them

Eigenvalue	λ_i	$\lambda_i / \lambda_T \cdot 100\%$
λ_1	2.214	82
λ_2	0.397	15
λ_3	0.060	2

For each stem the parameters were calculated according to equation 1 (b_1, b_2, b_3). Figs. 3, 4 and 5 show the geometrical effect of these parameters.

Parameter b_1 expresses the change in stem shape vertically onto the vertical axle of the stem. The dimension of this change is largest at the stem base and gradually decreases towards the top. Direction of the change is identical for all parts of the stem. Values of parameter b_1 for the individual Konšel's (Kraft's) tree classes are given in Table 2. It is obvious from the table that the arithmetical mean of this parameter increases from dominant trees to shady viable ones. Parameter b_1 is normally distributed (see Table 3). Its maximal value belongs to sample No. 34 which is represented by the shady viable tree 17 m long with d.b.h. 12.4 cm, tree overbark volume is 0.1106 m^3, crown volume 10.2 m^3. Minimal value belongs to sample No. 500. This tree can be included among dominant trees, its length is 23.5 m, d.b.h. 28.2 cm, tree overbark volume 0.8385 m^3, crown volume is 98.4 m^3.

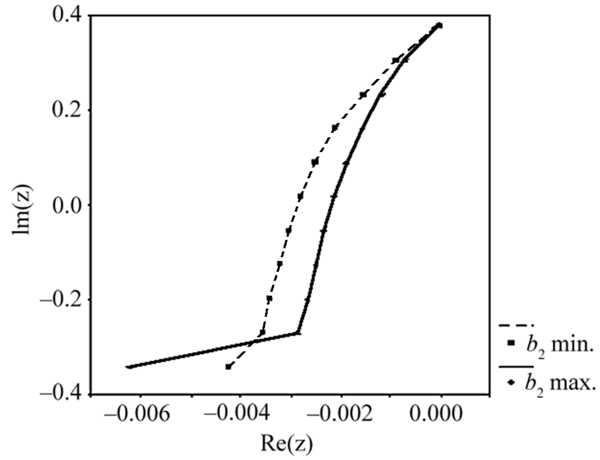

Fig. 4. Effect of changes in the second parameter

Parameter b_2 differs from the parameter b_1 in expression of shape change that has an opposite direction at the stem base and at the other parts of the stem. There are trees with great buttresses or individuals damaged by rot on the one hand, and on the other hand individuals with small buttresses. The integrated tendency of increase or decrease cannot be defined in relation to tree classes from Table 2. This parameter also has the normal distribution (see Table 3). Maximal value belongs to sample No. 29.

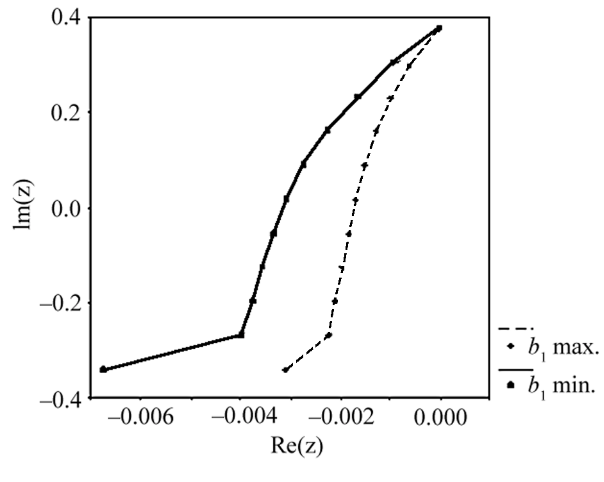

Fig. 3. Effect of changes in the first parameter

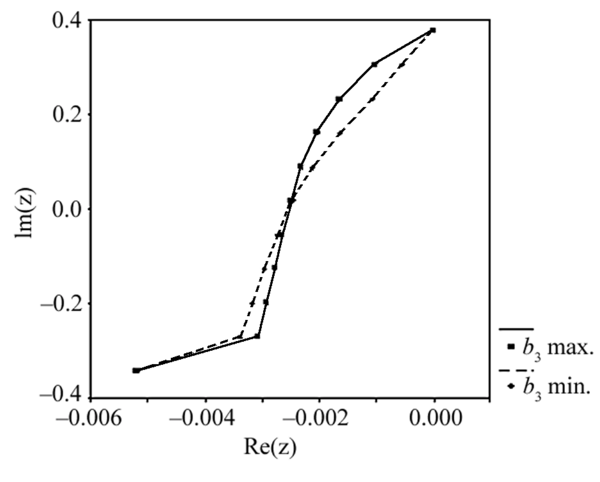

Fig. 5. Effects of changes in the third parameter

Table 2. Mean values of parameters b_1, b_2, b_3 for Konšel's tree classes (KONŠEL 1931), the corresponding Kraft's tree classification is presented in brackets

Tree class		b_1	b_2	b_3
	N	Mean . 10^{-3}	Mean . 10^{-4}	Mean . 10^{-5}
1 (1)	108	–1.53	–1.33	–7.98
2a (2)	225	–0.07	0.12	1.48
2b (3)	58	1.01	1.02	–2.41
3 (4a,b)	63	1.28	1.57	6.01
4 (5a)	27	1.56	–1.52	10.7

It is a shady viable tree, diameter at the stem base is 31 cm, d.b.h. 14.1 cm. On the contrary sample No. 380 is a dominant tree, diameter at the stem base 41 cm, d.b.h. 34.4 cm. The former sample shows large difference in diameters while the latter small difference in diameters.

The shape change expressed by the parameter b_3 is of the opposite direction in the lower and upper half of the stem with maximum at 8/10 of stem height. As regards the tree classes, no integrated tendency of increase or decrease is possible to define. This parameter does not have the normal distribution (see Table 3). Sample No. 338 shows the maximal value. It is a shady viable tree with alternate top. Further maximal values can be found in trees with dry or broken top. No dendrometric difference from the others was observed in trees with minimal values. It may be the trees with lengths measured in a wrong way or diameters in the top part of the stem.

In morphometry the dependence between shape and centroid size of the individual is often investigated (BOOKSTEIN 1991). The *configuration matrix* $\boldsymbol{X}$ is the $k \times m$ matrix of Cartesian co-ordinates of k landmarks in m dimensions. The *centroid size* is given by

$$S(\boldsymbol{X}) = \sqrt{\sum_{i=1}^{k} \sum_{j=1}^{m} (X_{ij} - \overline{X}_j)^2}, \ \boldsymbol{X} \in \boldsymbol{R}^{km}$$

where X_{ij} is the (i, j)th entry of $\boldsymbol{X}$, the arithmetic mean of the jth dimension is

$$\overline{X}_j = 1/k \sum_{i=1}^{k} X_{ij}$$

(DRYDEN, MARDIA 1998).

The hypothesis that the samples are based on the normal distribution in the case of centroid size and parameter b_3 can be rejected on the significance level $\alpha = 0.05$ (see Table 3).

Table 3. Kolmogorov-Smirnov tests of normality, DF – degrees of freedom, SIG. – significance level

	Statistic	DF	SIG.
b_1	0.037	481	0.158
b_2	0.036	481	0.170
b_3	0.041	481	0.048
S	0.080	481	0.000

Table 4. Tests of selected correlation coefficient (r) and Spearman coefficient of correlation (rs), for dependence between centroid size, parameters b_1, b_2 and b_3, sig. – significance level

	S	b_1	b_2
S	X		
b_1	rs = 0.248		
	sig. = 0.000	X	
b_2	rs = 0.088	r = 0.000	
	sig. = 0.055	sig. = 1.000	X
b_3	rs = 0.049	rs = 0.035	rs = 0.015
	sig. = 0.281	sig. = 0.438	sig. = 0.741

Table 4 contains the value of selected correlation coefficient and values of Spearman coefficient of correlation for quantities that do not have the normal distribution HAVRÁNEK (1993). The table also shows the values of significance for H_O: the quantities presented are non-correlated ones. Zero hypothesis can be rejected on the significance level $\alpha = 0.05$ only for the dependence between size and parameter b_1.

COMPLETE STEM ANALYSES

The assessment of stem shape development acquired from the complete stem analyses is an interesting possibility of PDM use. Four complete analyses of spruce stems were made on the plot Doubravčice 3 closely neighbouring with plot Doubravčice 1. Samples No. 171, 299, 301 and 313 were processed. Figs. 6 and 7 demonstrate the courses of the full Procrustes co-ordinates of samples No. 299 and 313. The age of the complete analyses is graded by five years.

The special PDM (inside-bark diameters) was elaborated for these samples.

It is evident from Fig. 8 that sample No. 313, which was an intermediate declining tree, is shifted in parameter values b_1 from the other three co-dominant samples. An increase in the values of parameter b_2 for sample No. 299 in Fig. 8 corresponds with the widening of the lower stem part in Fig. 7.

CONCLUSION AND DISCUSSION

The shape of spruce stem can be simplified into three sections (principal components) by using the principal components analysis. These sections are reflected in three parameters of the point distribution form model. The first PC explains 82% of variability on the experimental plot Doubravčice 1. The values of model b_1 connected with it show an increasing tendency from dominant trees to shady viable trees. This parameter has the normal distribution of frequencies. The second PC explains 15% of variability. With regard to Konšel's tree classes the integrated tendency of increase or decrease in the parameter b_2 is not possible to define. It also has the normal distri-

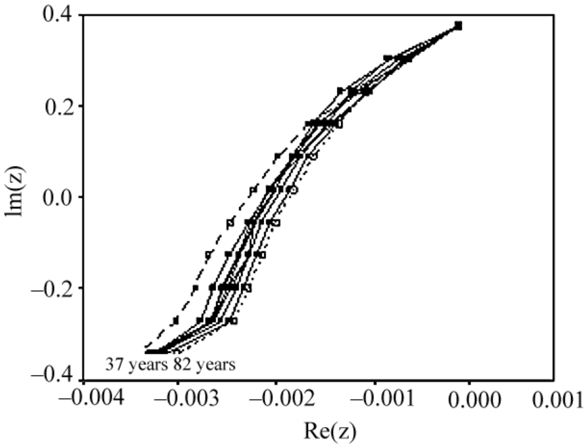

Fig. 6. Complete analysis of sample No. 313

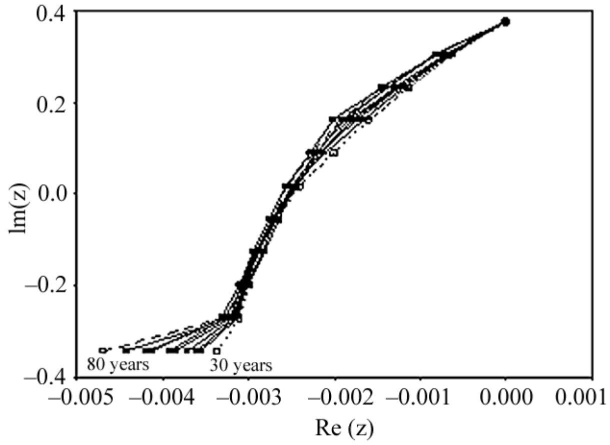

Fig. 7. Complete analysis of sample No. 299

bution of frequencies that is connected with distension of the lower stem part. The third PC explains 2% of variability. With regard to Konšel's classes the integrated tendency of increase or decrease in the parameter b_3 cannot be defined. It does not have the normal distribution of frequencies, which is connected with top breaks, top drying and also with faults during length measuring. If we identify the stem shape by the first PC (82% of variability), it is interesting that the course of the parameter b_1 is not dependent on age within the complete analyses, but rather on the competitive pressure on a tree. However, only 4 samples are taken into account and generalization of this conclusion will require more measurements.

The shape model is constructed for two dimensions. It is based on two vertically measured diameters and then on their mean values. A three-dimensional model could take into account other influences that certainly affect the spruce stem (cardinal points, terrain declination, prevailing wind direction, etc.). The method of deriving the configuration matrix is not the only possible way. For example, inflexion points could be added to the morphological curve, and on the contrary, the landmarks whose coordinates have the high value of correlation coefficient could be omitted.

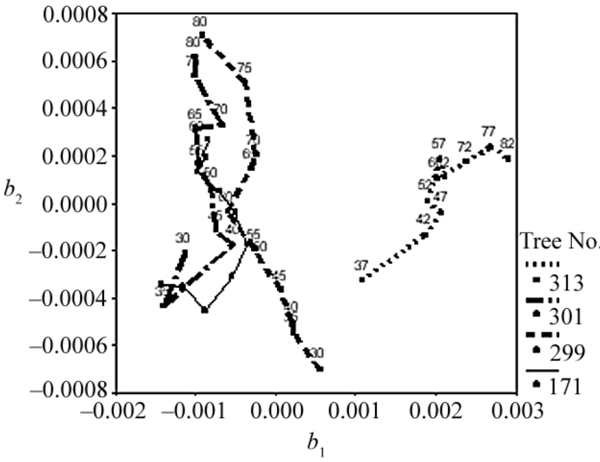

Fig. 8. Course of parameters b_1 and b_2 for the complete analysis of samples No. 171, 299, 301 and 313

References

BOOKSTEIN F.L., 1991. Morphometric Tools for Landmark Data. Cambridge, New York, Port Chester, Melbourne, Sydney, Cambridge University Press: 435.

COOTES T.F., TAYLOR C.J., COOPER D.H., GRAHAM J., 1992. Training models of shape from sets of examples. In: HOGG D.C., BOYLE R.D. (eds.), British Machine Vision Conference. Berlin, Springer Verlag: 9–18.

COOTES T.F., TAYLOR C.J., COOPER D.H., GRAHAM J., 1994. Image search using flexible shape models generated from sets of examples. In: MARDIA K.V. (ed.), Statistics and Images: Vol. 2. Oxford, Carfax: 1–18.

DRYDEN I.L., MARDIA K.V., 1998. Statistical Shape Analysis. Chichester, New York, Weinheim, Brisbane, Singapore, Toronto, John Wiley & Sons: 345.

HAVRÁNEK T., 1993. Statistika pro biologické a lékařské vědy. Praha, Academia: 478.

KONŠEL J., 1931. Stručný nástin tvorby a pěstění lesů v biologickém ponětí. Písek: 552.

KŘEPELA M., SEQUENS J., ZAHRADNÍK D., 2001. Dendrometric evaluation of stand structure and stem forms on Norway spruce (*Picea abies* [L.] Karst.) sample plots Doubravčice 1, 2, 3. J. For. Sci., *47*: 419–427.

KORF V. et al., 1972. Dendrometrie. Praha, SZN: 371.

METRIGUARD INC., 2001. www.metriguard.com.

ROHLF J.F., 1998. TpsRegr Version 1.20, Statistic software.

ROHLF J.F., 2001. http://life.bio.sunysb.edu/morph.

THOMPSON D.W., 1917. On Growth and Form. Cambridge, Cambridge University Press: 280.

Received 15 January 2002

Tvarový model vycházející z hraničních bodů pro smrkové kmeny (*Picea abies* [L.] Karst.)

M. KŘEPELA

Česká zemědělská univerzita, Lesnická fakulta, Praha, Česká republika

ABSTRAKT: Článek se zabývá konstrukcí tvarového modelu pro smrkové kmeny. Tento model vychází z analýzy hlavních komponent variančně-kovarianční matice sestavené pro Procrustova rezidua. Na praktickém příkladu smrkové zkusné plochy předmýtního věku je demonstrován výpočet Procrustových souřadnic, hlavních komponent a je sestaven tvarový model pro první tři komponenty. Jednotlivé parametry modelu jsou zhodnoceny vzhledem ke Konšelovým (Kraftovým) stromovým třídám. Je otestována normalita jejich rozdělení, maxima a minima jsou demonstrována na konkrétních stromech. Dále je provedena plná kmenová tvarová analýza čtyř vzorníků. Pro tyto vzorníky je sestaven zvláštní tvarový model a průběh parametrů tohoto modelu je graficky vyjádřen.

Klíčová slova: smrk ztepilý (*Picea abies* [L.] Karst.); tvar kmene; Procrustova analýza; analýza hlavních komponent; tvarový model vycházející z hraničních bodů

Tvar kmene je definován jako geometrická informace o konfiguraci po jejím ortogonálním posunutí, otočení a přeškálování (obr. 1). Konfigurace je tvořena tloušťkovými a výškovými souřadnicemi jednotlivých hraničních bodů na morfologické křivce kmene. Hraničních bodů je celkem 21. Jsou umístěny na pařezu a dále postupují po 1/10 výšky kmene k jeho vrcholu (obr. 2). Ortogonální transformaci zabezpečují Procrustovy souřadnice. Procrustův plný tvarový průměr $\hat{\mu}$ je vypočítán jako vlastní vektor příslušející k největší vlastní hodnotě variančně-kovarianční matice, vypočtené pro pre-tvary.

Následně jsou vypočteny plné Procrustovy souřadnice, které minimalizují vzdálenost mezi tvarovým průměrem a jednotlivým tvarem.

Konstrukce vlastního tvarového modelu vychází z prací COOTESE et al. (1992, 1994).

Tvarový model je založen na Procrustových reziduích (odchylkách plných Procrustových souřadnic od plného Procrustova průměru). Z těchto reziduí je vypočtena variančně-kovarianční matice. Pro tuto matici je je provedena analýza hlavních komponent. Je vypočteno t vlastních hodnot a k nim odpovídající vlastní vektory.

Jako praktický příklad posloužil materiál ze zkusné plochy Doubravčice 1 na ŠLP Kostelec nad Černými lesy. Na této zkusné ploše bylo proměřeno celkem 481 kmenů. Bylo také provedeno jejich zatřídění do Konšelových (Kraftových) stromových tříd. Pro každý kmen byly sestaveny konfigurace, pro celý soubor byl vypočten plný Procrustův průměr a následně byly vypočteny plné Procrustovy souřadnice pro každý kmen. Pro účely odvození modelu byly vypočteny vlastní hodnoty a vlastní vektory pro variančně-kovarianční matici vypočtenou z Procrustových reziduí. Přehled prvních tří hlavních hodnot, které odrážejí 99 % celkové variability, je zachycen v tab. 1.

Pro každý kmen byly dále vypočteny parametry b_1, b_2 a b_3. Aritmetické průměry těchto parametrů byly vypočteny pro jednotlivé Konšelovy (Kraftovy) stromové třídy a jejich přehled udává tab. 2. Maxima a minima těchto vah jsou zachycena na obr. 3–5.

Parametr b_1 má normální rozdělení (tab. 3) a roste od „neplnodřevných" (předrůstavých) stromů po „plnodřevné" (zastíněné životaschopné) stromy.

Parametr b_2 odráží zbytnění oddenku kmene. Má normální rozdělení (tab. 3). Minimální hodnota je u zastíněných životaschopných stromů, dále u předrůstavých stromů. Potom následují stromy úrovňové hlavní, vedlejší a vrůstavé ustupující.

Parametr b_3 odráží stromy s defektním vrcholem a stromy se špatně změřenou délkou. Nemá normální rozdělení (tab. 3).

Dále byl zkoumán vztah mezi centrální velikostí kmene a jednotlivými parametry. Tato závislost byla prokázána pouze mezi centrální velikostí a parametrem b_1 (tab. 4), tedy čím větší centrální velikost, tím je strom méně plnodřevný.

Následně byla provedena tvarová analýza pro čtyři kmeny, které byly podrobeny plné kmenové analýze. Jednalo se o tři úrovňové hlavní stromy a jeden vrůstavý ustupující strom. Pro tyto kmeny byly vypočteny plné Procrustovy souřadnice a sestaven tvarový model. Obr. 6 a 7 zachycují tvary získané z plných kmenových analýz po pětiletém intervalu. Parametry b_1, b_2 modelu jsou zachyceny na obr. 8. Zde je nápadný posun vrůstavého ustupujícího stromu od zbývajících tří úrovňových hlavních stromů.

Corresponding author:

Mgr. Ing. MICHAL KŘEPELA, Česká zemědělská univerzita, Lesnická fakulta, katedra hospodářské úpravy lesů, 165 21 Praha 6-Suchdol, Česká republika
tel.: + 420 2 24 38 37 18, fax: + 420 2 20 92 31 32, e-mail: krepela@lf.czu.cz