P S Z C Z E L N I C Z E Z E S Z Y T Y N A U K O W E
Rok XLIV, Nr 2 2000

A METHOD OF CLASSIFYING HONEY BEE RACES BY THEIR BODY CHARACTRS

K. Čermák, F. Kašpar*

Bee Research Institute Ltd., Dol, Breeding Stn. Zubří, 756 54 Zubří, Czech Republic
* Bee Research Institute Ltd., Dol, Breeding Stn. Pekařov, 788 23 Jindřichov, Czech Republic

S u m m a r y

A method to classify honey bee races based on morphological characters of bee-worker forewings was developed. Fifteen to 20 wings from a colony are scanned as bitmap pictures. The co-ordinates of 19 defined points in wing veins are measured using the mouse pointer, and 30 characters (lengths, angles, indices, fields area) are çomputed. The combination of values of characters is typical of definite race. The classified sample is compared with known values of single races. The Mahalanobis distances of the sample to single races are computed. The final result of the sample classification are posterior probabilities with which the sample belongs to single races. Classification results of some bee samples are given.

Keywords: honeybee worker, morpholigcal charactrs, races, Mahalanobis distances.

INTRODUCTION

Maintaining the biodiversity within the honey bee species assumes having adequately quick and accurate method of honey bee classification. Complex morphological analysis of bee exterior is labour intensive. The more simple method is measuring the bee forewing characters which are typical of each pure race (K a u h a u s e n - K e l l e r D . , K e l l e r R . , 1994).

The bee samples are usually evaluated using the discriminant analysis and are presented graphically for the whole populations - races. But the identification and classification of a sample is difficult. At the same time, expressing the multidimensional relations among single populations in a two-dimensional diagram constricts the informational ability of the diagram.

Using the principles of discriminant analysis we derived a method of single bee sample classification in numerical form. A sample is classified by the probability of belonging to a single race, which is characterised by race standards prepared in advance.

MATERIAL AND METHODS

The measurement of bee wing characters is easy. Forewings are fixed on a transparent foil. Wings are scanned at the resolution at least 1200 dpi and moved into computer as a bitmap picture. The wing design is cut in a program on the screen with a mouse pointer so that x,y co-ordinates of given points in given order are saved. 30 characters are then computed on the basis of co-ordinates of single points for each wing. These characters determine the design of the whole wing. The wing characters are lengths between some points, angles, indices and area of six wing fields. 15 - 20 wings in one sample are measured. A mean is computed for each character as a parameter of the sample.

The sample classification is made by comparing with the typical values of single pure races - standards prepared beforehand. Bee samples for these standards were collected from various beekeepers in Europe, Asia and Africa. Race standards were repeatedly statistically checked by excluding samples which diverged from other typical samples of the race.

The match between the samples, tested and standard, in terms of race-characteristic values is expressed in ***Mahalanobis distances***. These are distances of the sample from centroids of clusters of single races. Mahalanobis distances are calculated by the formula (Lukasová A., Šarmanová A., 1985):

$$MD_{a,b} = (a - b)^T \times K^{-1} \times (a - b)$$

where:

$MD_{a,b}$ ..	Mahalanobis distance between points *a,b* in multidimensional space
$(a-b)^T$ ...	line vector $(a_1-b_1, a_2-b_2, ..., a_n-b_n)$ transposed
K^{-1} ...	inverse matrix to covariance matrix of standardised data
(a-b) ...	column vector $(a_1-b_1, a_2-b_2, ..., a_n-b_n)$

A sample belongs with the highest probability to that race which it is most close to, i.e. has the lowest Mahalanobis distance to it. The probability is computed for each race whose standard is known. The sum of all probabilities for a sample is 100. These are so called ***posterior probabilities***. They are computed from the values of Mahalanobis distances.

Posterior probability (PP) that a sample *t* belongs to population *i* is

$$PP(t) = \frac{\frac{1}{n} * \exp(-MD(i)^2 / 2)}{\sum_i \frac{1}{n} * \exp(-MD(i)^2 / 2)}$$

where:

t ...	tested sample
i ...	single populations in the set of standards
n ...	number of populations in the set of standards
$MD(i)^2$ ...	the 2nd power of Mahalanobis distance of the sample *t* from the population *i*
exp() ...	the function calculating the value of e^x, where *x* is given number,
e is a	base of natural logarithms - around 2,71828

RESULTS AND DISCUSSION

The statistical processing of single wings and classification of the whole sample runs in database application of FoxPro named BEEMORPH. An example of resulted values of Mahalanobis distances, classifications to races and corresponding posterior probabilities of a few samples are in the Table 1. Samples were compared with nine race standards. The values of the first five races by the order of posterior probabilities are given in the Table 1. Samples C0413 and C0477 were classified to different races than had been declared.

Tab. 1

Classification results of some samples to first five of nine race standards
Zakwalifikowanie próbek na podstawie uzyskanych wyników do pięciu ras spośród dziewięciu porównywanych modeli

Sample No.	Declared race		Classification of the sample at				
			1^{st} place	2^{nd} place	3^{rd} place	4^{th} place	5^{th} place
C0022	carnica	*MD*	4,890968	5,574902	5,603711	5,728217	5,997851
		classified race	carnica	macedonica	caucasica	mellifera	ligustica
		PP (%)	93,76	2,62	2,23	1,10	0,23
C0034	carnica	*MD*	7,248234	7,301153	7,330718	7,547787	7,618142
		classified race	carnica	buckfast	mellifera	caucasica	macedonica
		PP (%)	40,61	28,31	22,81	4,54	2,66
C0087	buckfast	*MD*	5,242633	5,744137	5,824023	5,927215	5,968088
		classified race	buckfast	ligustica	carnica	mellifera	macedonica
		PP (%)	86,58	5,51	3,47	1,89	1,48
C0274	iberica	*MD*	4,901776	6,485338	6,620067	6,664406	6,702301
		classified race	iberica	carnica	buckfast	mellifera	macedonica
		PP (%)	99,97	0,01	0,01	0,00	0,00
C0413	mellifera	*MD*	5,206643	5,279640	5,379239	5,481054	5,629261
		classified race	carnica	mellifera	ligustica	caucasica	macedonica
		PP (%)	39,74	27,10	15,94	9,17	4,03
C0464	ligustica	*MD*	4,260699	4,712279	5,034632	5,068578	5,093799
		classified race	ligustica	carnica	caucasica	mellifera	macedonica
		PP (%)	82,79	10,92	2,27	1,91	1,68
C0477	cecropia	*MD*	7,072516	7,274367	7,328876	7,648411	7,675427
		classified race	macedonica	cecropia	carnica	mellifera	ligustica
		PP (%)	70,18	16,50	11,08	1,01	0,82

MD ... Mahalanobis distance
PP ... posterior probability

Should the value of posterior probabilities for a race in the first place be lower than 50%, the sample is not reliably compared to any race. At the same time, samples in the first two or three places with similar values of posterior probabilities may be considered as samples with the current impact of a few races (samples C0034, C0413).

Only those samples were left as typical which were classified identically with declared race and with the value of posterior probability at the first place over 70%, during creating the race standards. This criterion fits in with samples C0022, C0087, C0274, C0464 in Table 1.

The typical values of 30 characters of four races (standards) in relation to the mean of nine races are graphically presented in Fig. 1. The combination of characters is specific for each race. These combinations may be considered to be race standards. Correspondingly, classification of the single sample using the above described multivariate method is made by comparing it to standards and expresses the similarity of the sample to the race standards.

The described method of the morphometric analysis and the taxonomic classification of honey bees named DAWINO-bee has been employed in checking the race purity in Carniolan bee breeding apiaries in the Czech Republic since 1999. The computer program is able to print a protocol with detailed results of the sample. An example of the main part of the protocol - result of the sample classification is in Tab. 2.

Tab. 2

Results of the taxonomic classification of a honey bee sample by wing characters - Wyniki klasyfikacji taksonomicznej próbek pszczół na podstawie morfologicznych cech skrzydła

RACE	PP %
CARNICA	92
MACEDONICA	4
MELLIFERA	2
LIGUSTICA	1
Other	1

PP ... posterior probability

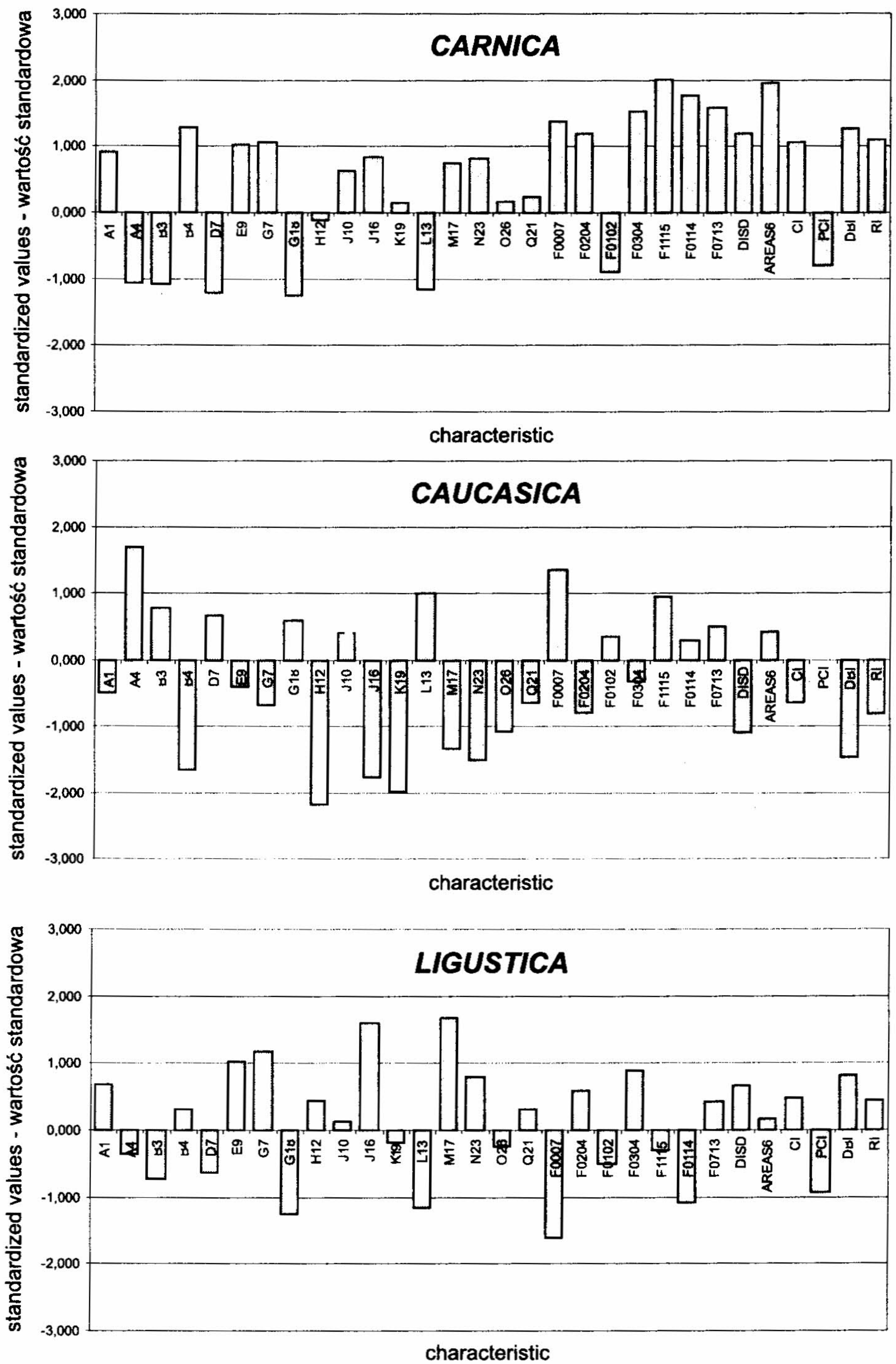

Fig. 1 Combination of 30 characters of four bee races expressed by deviations from a mean of nine races - Kombinacja 30 cech czterech ras pszczół przedstawiona poprzez odchylenie od średnich dziewięciu ras

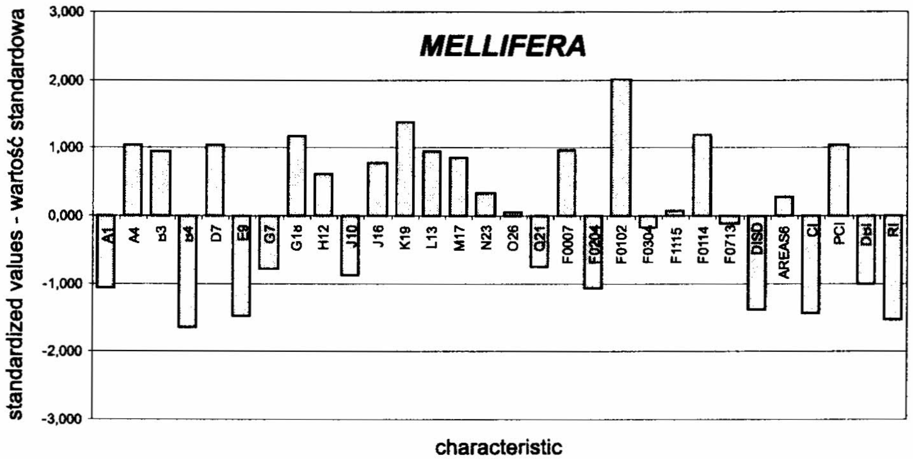

Fig. 1 Combination of 30 characters of four bee races expressed by deviations from a mean of nine races - Kombinacja 30 cech czterech ras pszczół przedstawiona poprzez odchylenie od średnich dziewięciu ras

REFERENCES

Kauhausen-Keller,D., Keller,R. (1994)- Morphometrical control of pure race breeding in honeybee (*Apis mellifera* L). *Apidologie* 25, 133-143.

Lukasová, A., Šarmanová, A. (1985)- Metody shlukové analýzy. *SNTL* Praha.

OCENA PRZYNALEŻNOŚCI RASOWEJ PSZCZÓŁ NA PODSTAWIE CECH MORFOLOGICZNYCH

Čermák K., Kašpar F.

Streszczenie

Opracowano metodę oceny przynależności rasowej pszczół na podstawie pomiaru cech na skrzydle I pary pszczół robotnic.

Pomiary wykonano na skrzydłach 15-20 pszczół z każdej rodziny. Na zeskanowanym obrazie skrzydła zaznaczono na monitorze komputera 19 punktów, z których program komputerowy wyliczał 30 cech (odległości punktów, kąty, powierzchnie). Wartości te porównywano następnie z wcześniej przyjętymi modelami rasowymi co pozwoliło na określenie przynależności rasowej pszczół z badanej rodziny. Przestawiono wyniki pomiarów niektórych próbek.

Słowa kluczowe: pszczoła robotnica, cechy morfologiczne, analiza komputerowa, przynależność rasowa.