TÜBİTAK

Turkish Journal of Agriculture and Forestry

http://journals.tubitak.gov.tr/agriculture/

Research Article

Turk J Agric For
(2014) 38: 111-124

doi:10.3906/tar-1305-8

Artificial neural networks in online semiautomated pest discriminability: an applied case with 2 *Thrips* species

Peter FEDOR[1,*], Eladia Maria PEÑA-MÉNDEZ[2], Halina KUCHARCZYK[3], Jaromír VAŇHARA[4], Josef HAVEL[5], Martina DORIČOVÁ[1], Pavol PROKOP[6]

[1]Department of Environmental Ecology, Faculty of Natural Sciences, Comenius University, Bratislava, Slovakia
[2]Department of Analytical Chemistry, Nutrition, and Food Science, Faculty of Chemistry, University of La Laguna, Campus de Anchieta, La Laguna, Tenerife, Spain
[3]Department of Zoology, Faculty of Biology and Biotechnology, Maria Curie-Sklodowska University, Lublin, Poland
[4]Department of Botany and Zoology, Faculty of Science, Masaryk University, Brno, Czech Republic
[5]Department of Chemistry, Faculty of Science, Masaryk University, Brno, Czech Republic
[6]Department of Biology, Faculty of Education, Trnava University, Trnava, Slovakia

Received: 03.05.2013 • **Accepted:** 09.07.2013 • **Published Online:** 13.12.2013 • **Printed:** 20.01.2014

Abstract: Being faced with practical problems in pest identification, we present a methodical paper based on artificial neural networks to discriminate morphologically very similar species, *Thrips sambuci* Heeger, 1854 and *Thrips fuscipennis* Haliday, 1836 (Thysanoptera: Thripinae), as an applied case for more general use. The artificially intelligent system may be successfully applied as a credible, online, semiautomated identification tool that extracts hidden information from noisy data, even when the standard characters have much overlap and the common morphological keys hint at the practical problem of high morphological plasticity. Statistical analysis of 17 characters, measured or determined for each *Thrips fuscipennis* and *T. sambuci* specimen (reared from larvae in our laboratories), including 15 quantitative morphometric variables, was performed to elucidate morphological plasticity, detect eventual outliers, and visualize differences between the studied taxa. The computational strategy applied in this study includes a set of statistical tools (factor analysis, correlation analysis, principal component analysis, and linear discriminant analysis) followed by the application of a multilayer perceptron artificial neural network system, which models functions of almost arbitrary complexity. This complex approach has proven the existence of 2 separate species: *T. fuscipennis* and *T. sambuci*. All the specimens could be clearly distinguished with 2 distinct subgroups for each species, determined by sex. In conclusion, the use of an optimal 3-layer ANN architecture (17, 4, 1) enables fast and reliable 100% classification as proven during the extensive verification process.

Key words: Artificial neural networks, online semiautomated pest identification, Thysanoptera

1. Introduction

Phytosanitary field technical staff often face the risk of crop damage, which has increased recently with the introduction of new (sometimes exotic) species to plant material. Correct identification of the targeted pest insects is essential for phytosanitary management. Mistakes at this stage can cause project delays or failure.

Like in any insect group, thrips (Thysanoptera) identification includes a wide range of alternatives and specific methods (Mehle and Trdan, 2012), from printed dichotomous taxonomic keys (e.g., Schliephake and Klimt, 1979; zur Strassen, 2003 for European species) to more user-friendly pictorials (Mound and Kibby, 1998) and complex multiaccess keys (Moritz et al., 2001). A computerized knowledge base, using HyperWriter (NTERGAID, Fairfield, CT, USA), was developed to enable vegetable producers, field technical staff, extension personnel, and other nonentomologists to identify the species of thrips on economically important thrips-infested vegetable crops (Frantz and Mellinger, 1997). For 15 years, Lucid, a digital matrix key system, has been evolving to keep pace with technological advances in software (Schuetz et al., 2010; Taylor, 2010). Genetic markers have also proven to be a powerful tool in the identification of thrips pest species, including their immature stages (Moritz et al., 2000; Brunner et al., 2002; Toda and Komazaki, 2002; Rugman-Jones et al., 2006). The interactive electronic key created by Moritz et al. (2004) combines both morphological and molecular information.

The latest review comparing "traditional" and "modern" methods in thrips identification (Mehle and Trdan, 2012) summarized all of the alternatives mentioned

* Correspondence: fedor@fns.uniba.sk

above, even gently annotating the power of morphometrics in encompassing size and shape of a biological object. All species exhibit morphological variation, induced both genetically and by the environment as a phenotypic plasticity (Ananthakrishnan, 2005), which can serve as a buffering mechanism against environmental changes. Subsequently, many species have adults so varied in structure that large and small individuals may not be recognizable as being the same species without collateral biological information (Mound, 2005) or brand new identification keys (if available). However, our previous studies (Fedor et al., 2008, 2009; Kucharczyk and Kucharczyk, 2009; Kucharczyk et al., 2012) indeed underline morphometric (both quantitative and qualitative) variables as an autonomous tool for reliable Thysanoptera identification, in the sense of either basic multivariable analyses (e.g., principal component analysis) or even the phenomenon of more complex artificial neural networks (ANNs). Variability, in fact, may sometimes appear rather subjective, sometimes even objectively proven by basic statistical tools analyzing data within their interval range. However, characters, although widely dispersed, together often form a unique set with original patterns hardly recognizable within a simple statistical approach. ANNs, when applied for a set of carefully selected morphometric variables, may provide a solution to determining a species (Vaňhara et al., 2007; Muráriková et al., 2011). Of the previously analyzed characters (Fedor et al., 2008, 2009), they were all (head width and length, eye length, ovipositor length, antenna length, and distance between an anterior and posterior pair of ocelli) used in many traditional keys (e.g., Schliephake and Klimt, 1979; zur Strassen, 2003 for European species), but, however, only as additional data. Artificially intelligent systems, at least in specific cases, offer a way to evaluate them as an autonomous set of data.

Undoubtedly, in the recent decades, there has been a growing interest in ANN systems, which, in fact, have many forms and versions; however, in general they have 2 important factors in common: ability to learn from examples and to generalize the observed patterns (Weeks and Gaston, 1997; Gaston and O'Neill, 2004). Although the ANN systems correspond to the theory of how real biological neurons (neuronal networks) process received information and there are many elementary similarities between the human brain and artificial intelligence (e.g., learning from experience and storing information as patterns), the synaptic connections in artificial networks analyze both positive and negative values and are often implemented to evaluate data out of the neurobiological background (Freeman and Skapura, 1992; Ripley, 1993; Haykin, 1994; Haralabous and Georgakarakos, 1996).

ANN models are flexible function approximators to describe nonlinear systems (Zhang and Barrion, 2006), make no a priori assumptions on the type or statistical distribution of data, and, thus, can be used for pattern recognition on practically any kind of multivariate data sets (Do et al., 1999; Moore and Miller, 2002; Clark, 2003; Kavdır, 2004; Marini et al., 2004; Aldrich et al., 2007; Vaňhara et al., 2007; Fedor et al., 2008, 2009; Esteban et al., 2009; Muráriková et al., 2011; Bilgili, 2011; Tohidi et al., 2012). Their use now spans all fields of science, including a wide variety of applied branches, such as pest control in agriculture and forestry.

Being encouraged by the previous research, which has undisputedly built up a theoretical (mathematical) background for the opportunity of ANN insect species identification (Vaňhara et al., 2007; Fedor et al., 2008, 2009; Muráriková et al., 2011), we describe a methodical concept and the power of artificially intelligent systems in discriminating 2 morphologically similar *Thrips* species as a relatively simple real model case for applied entomology (pest identification). This is, in fact, the first time that artificial intelligence has been applied for 2 very similar and often hardly recognizable (at least by technical staff) pest species with overlapping morphometric characters and limited material. We present the power of morphometrics as an autonomous set of information for reliable species discrimination with applied proposals in the final computational products (semiautomatic online identification system).

As an applied model, discriminability of *Thrips sambuci* Heeger, 1854 and *T. fuscipennis* Haliday, 1836 (Thysanoptera: Thripinae) has been studied. *Thrips* Linnaeus, 1758 is the most species-rich genus of Thysanoptera with more than 250 described species worldwide (Mound, 2010). There are consequentially many identification systems available (e.g., Nakahara, 1994; zur Strassen, 2003; Mound and Masumoto, 2005; Mound and Azidah, 2009; Mound, 2010; Vierbergen et al., 2010). First and second larval instars of Thripidae differ in the number of setae on the pronotum (6 and 7 setae). Generally, they are easy to recognize by 1 (first instar larva) or 3 (second instar larva) pairs of setae on abdominal sternites IV–VIII and 3–4 (first instar larva) or 4–5 (second instar larva) pairs of setae on abdominal segment IX posteromarginally (Kucharczyk, 2010; Vierbergen et al., 2010).

However, in applied phytosanitary pest monitoring, problems with prompt and clear species identification may often occur, for instance, if visible differences between similar species are none or minute. Although the ongoing development of molecular technology and computational strategy offers tools of ever-increasing speed, there are still several species that are more difficult to determine. Within Thripinae, for instance, there are sometimes practical problems and confusion in distinguishing between *T. fuscipennis* and *T. sambuci* adults, which is in contradiction with their larvae, who are morphologically

significantly different and thus easy to identify (Kucharczyk, 2010; Vierbergen et al., 2010; Kucharczyk et al., 2012). While European *Thrips sambuci* prefers shrubs *Sambucus nigra* L. and *S. racemosa* L., the very similar *T. fuscipennis* (introduced to Asia and North America) is a pest in temperate greenhouses (Jacobson, 1995) predominantly causing damage to cucumber (*Cucumis sativus* L.) foliage (silvering, bronzing) or flowers. Even ornamentals, such as roses (*Rosa* spp.), may have leaves silvered and petals covered by brown patches under their infestation (Henneberry et al., 1961). In Central Europe, both populations often occur in parallel, particularly after *T. sambuci* infiltrates farmland when overpopulated in ecotonal shrub stands.

Unfortunately, although being different morphologically (zur Strassen, 2003; Kucharczyk et al., 2012), especially in chaetotaxy and color of antennal segments, the identification of many specimens that may have overlapping characters may be sometimes confusing, especially for technical staff with less taxonomic experience; if, for instance, a sample is extracted from sticky traps (damaged material); or if mounted material of poor quality is available for an expert. Moreover, although the traditional morphological methods for the identification of thrips pests have been recently accompanied by DNA (Brunner et al., 2002) or protein analysis (e.g., Toda and Komazaki, 2002) to make the control more reliable and dynamic (Mehle and Trdan, 2012), the BLAST (FN546130.1 and FN546131.1, FN546127.1 and FN546129.1) discrimination of *Thrips sambuci* and *T. fuscipennis*, when financially accessible, still remains limited due to the selected mitochondrial gene sequences and requires revision in GenBank. For such cases, ANN systems may be usefully applied as an occasional and specific alternative.

2. Materials and methods

2.1. Materials

In total, 175 *Thrips* specimens of 2 species, *T. fuscipennis* and *T. sambuci*, were collected for their detailed morphometric revision within statistical and ANN analyses or used for verification of the online discrimination system. The computational set (matrix) consisted of 93 specimens of both evaluated species, *T. fuscipennis* (27 females and 15 males) and *Thrips sambuci* (36 females and 15 males), and the verification set of 82 more specimens that originated from the larger Central and East European area (Poland, Slovakia, Austria, and the Czech Republic) to record wider intraspecific morphological variability and to prove the reliability of the system. They were sampled mainly from farmland on *Pisum sativum* monocultures (*T. fuscipennis*) with *Sambucus nigra* solitaires (*T. sambuci*). In order to ensure a more reliable database, most of the *T. sambuci* specimens analyzed in Poland were kept for the duration of their life cycle from eggs to adults (larvae are easily determined) in Fytotron plant growth chambers (photophase: 16 h at 24 °C, scotophase: 8 h at 10 °C, on *Sambucus nigra* as the host plant) in the laboratory (Department of Zoology Laboratory of Maria Curie-Sklodowska University, Lublin, Poland). The material from Slovakia, Austria, and the Czech Republic comes from older permanent slides stored in our collections (Comenius University, Bratislava, Slovakia), sometimes with no additional information on host plants. Standard preparatory technique was used for mounting; specimens were collected into AGA (a mixture of ethyl alcohol, glycerin, and acetic acid), macerated in warm 10% KOH, dehydrated in alcohol and clove oil, and mounted on slides in Canada balsam. The material was identified by P Fedor and H Kucharczyk and is deposited in their collections.

2.2. Selection of characters

We defined and recorded a total of 17 characters (Table 1, Figure 1), measured for each specimen of both species (*Thrips sambuci* and *T. fuscipennis*) collected. Sixteen of them may be defined as quantitative morphometric (15) or qualitative (1 - number of campaniform sensilla on mesonotum) traits related to different parts of the body, including the head (1–7), thorax (8–13), and abdomen (14–16). Most of the selected characters are commonly used for thrips identification (e.g., zur Strassen, 2003). The 17th variable indicates sex. The morphometric characters were measured quantitatively as linear distances on digital images taken from slide-mounted specimens by H Kucharczyk (microscope OLYMPUS BX 61 and image analyzer software sellSens Dimension Ver. 2010, Poland) and P Fedor (microscope LEICA M 205 C and image analyzer software LUCIA net, Laboratory Imaging Ltd., Slovakia and Czech Republic). Both species, like many other thrips, exhibit a pronounced sexual dimorphism (zur Strassen, 2003). The data corresponding to missing structures in males (ovipositor) and females (area porosae on sternum V–VI) were included in the analysis and entered into the data matrix as empty cells.

2.3. Software and computational strategy

ANN computation was performed using the TRAJAN Neural Network Simulator, Release 3.0 D. (TRAJAN Software Ltd. 1996–1998, UK) and the program STATISTICA 6 (StatSoft, Inc., Tulsa, OK, USA). All computations were performed on a standard PC computer with operating system Microsoft Windows Professional XP 2003 and/or MW 2010.

All the statistical methods applied to evaluate the set of morphometric data related to *T. sambuci* and *T. fuscipennis* are commonly (perhaps except for the ANN) used in taxonomy (e.g., Chiapella, 2000; Apuan et al., 2010), each, obviously, with its own specific approach. Therefore, there

Table 1. A statistical survey of the characters applied for the analysis (length in µm, uncertainty 0.03 µm; SD - standard deviation, V - variable; * - mean and standard deviation irrelevant). 1 - head width; 2 - head length (dorsal side); 3 - head length (ventral side, including mouthcone); 4 - eye length; 5 - antennal segment V length; 6 - antennal segment VI length; 7 - distance between an anterior and posterior pair of ocelli; 8 - distance between CS and metanotum; 9 - distance between D1 and metanotum; 10 - length of posteroangular seta interna; 11 - length of posteroangular seta externa; 12 - number of CS on mesonotum; 13 - distance between setae D1 and fore edge on metanotum; 14 - ovipositor length; 15 - width of area porosae on sternum V; 16 - width of area porosae on sternum VI; and 17 - sex.

V.	*T. fuscipennis* male		*T. fuscipennis* female		*T. sambuci* male		*T. sambuci* female	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
1	126.33	14.33	144.07	13.23	126.00	4.71	150.42	4.41
2	89.00	8.49	95.00	6.50	86.33	7.67	103.47	6.53
3	164.67	12.60	192.78	13.61	184.00	4.31	215.28	9.48
4	53.67	4.81	58.61	2.89	55.33	1.29	59.72	2.66
5	32.00	2.15	36.02	2.52	33.50	1.84	37.36	2.23
6	50.17	2.75	51.57	2.10	52.00	2.15	55.07	1.84
7	14.17	1.54	16.20	1.75	13.17	1.48	14.65	1.81
8	11.92	2.73	18.80	4.28	9.89	8.56	27.01	9.38
9	10.00	1.64	13.98	2.22	13.67	1.86	22.08	3.13
10	43.67	4.10	56.57	3.74	52.17	2.48	67.36	5.51
11	38.83	3.39	51.48	4.40	44.33	3.34	60.00	4.93
12	*	*	*	*	*	*	*	*
13	13.16	3.06	14.44	3.35	17.97	2.52	20.74	3.95
14	*	*	205.00	8.32	*	*	201.25	6.25
15	43.83	5.58	*	*	35.17	5.04	*	*
16	37.83	6.67	*	*	24.33	6.37	*	*
17	*	*	*	*	*	*	*	*

is no need for their detailed description. In this paper, the evaluation includes application of factor analysis, correlation analysis, principal component analysis, linear discriminant analysis, and ANN analysis.

The ANN computational strategy applied in this study was introduced in our previous studies (Vaňhara et al., 2007; Fedor et al., 2008). Data were randomly divided into a learning (training) set, a verification set, and a test set. Each set consisted of a number of samples (thrips specimens) characterized by input variables (characters) and identified as belonging to a species (output). Preliminary experiments suggested that the multilayer perceptrons (MLPs) would be the most efficient tool for this purpose. MLP is generally one of the most commonly used types of ANN and can model functions of almost arbitrary complexity. Its architecture is conventionally constructed with 3 or more feed-forward layers, i.e. input, output, and 1 or more hidden layers. Each layer might have a different number of nodes.

The learning or training process of MLP consists of searching for such values of w_{ij} weights to minimize the root mean square (RMS) value:

$$\mathrm{RMS} = \sqrt{\frac{\sum_{i=1}^{N}\sum_{j-1}^{M}(y_{ij} - out_{ij})^2}{N \times M}}$$

where y_{ij} is the element of the matrix (N × M) for the training set, out_{ij} is the element of the output matrix (N × M) of the neural network, N is the number of variables in the pattern, and M is the number of samples. By running the data on specimens from the training set, including the output variable (the identification), through the network and comparing the actual output generated with the desired or target outputs, the network automatically adjusts the weights and thresholds in order to minimize the overall error.

The training of a MLP network can be executed by different algorithms. We used back propagation, which

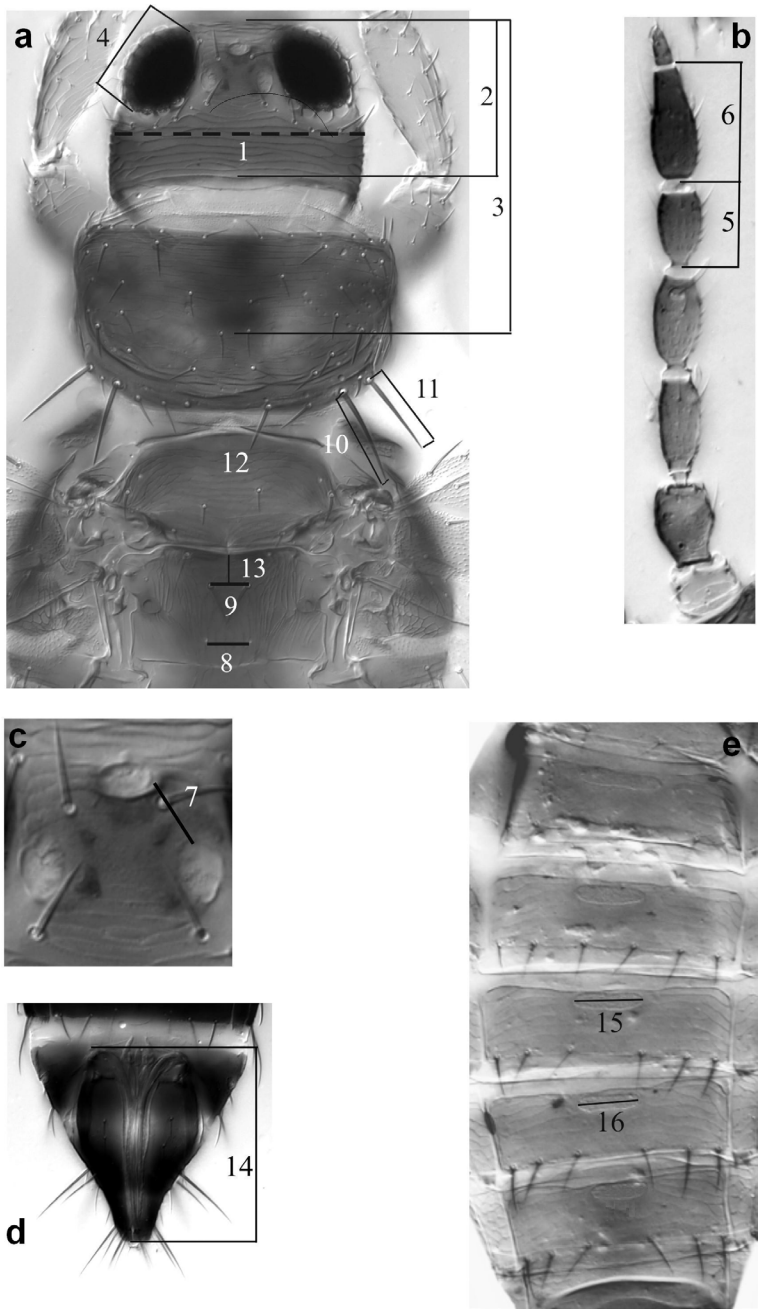

Figure 1. Morphological characters applied for *Thrips* identification (explanation in Table 1). a) head and thorax; b) antenna; c) ocelli; d) ovipositor; e) abdomen ventrally (a–d in *T. sambuci*; e in *T. fuscipennis*).

is the best-known one and has relatively low memory requirements (Fausett, 1994; Patterson, 1996). We ran the training algorithm several times with each configuration for 5000 to 10,000 iterations (epochs) to ensure a proper convergence to RMS minimum and to avoid being stuck in a local minimum. After obtaining the optimal architecture and minimal RMS, a number of randomly selected specimens from the learning set, as well as individuals from different populations, were excluded to form the verification set. The verification is a test of prediction power of the model.

ANN computation also comprises pre- and postprocessing stages. Preprocessing techniques used in our study included data standardization (scaling) and conversion of nominal input variables to numeric values (done automatically by TRAJAN software).

Similarly, the output activation scores were transformed by postprocessing into the name of thrips species. The classification by ANN is performed by checking output unit activation levels against 2 thresholds, the accept threshold and the reject threshold. To simply assign the classification to the species corresponding to the winning unit, irrespective of the settings of other units, we set the accept threshold to zero and ignored the reject threshold.

Following the trained neural network architecture, the online web application for semiautomated discrimination of both species, using PHP programming language, was developed and based on extraction of the TRAJAN net weights and the activation function.

3. Results

3.1. Basic statistical data analysis

Statistical analysis on 17 characters, measured or determined for each *Thrips fuscipennis* and *T. sambuci* specimen (Table 1), including 15 quantitative morphometric variables on different parts of their bodies, was performed to elucidate morphological plasticity, detect eventual outliers, and visualize differences between the studied taxa.

Basic evaluation of both species compared graphically (Figure 2), including their character average values as well as standard deviations (Table 1), calculated separately for males and females, hints at their high intraspecific biological variability, while interspecifically most of their average values appear too similar to distinguish the species reliably.

The number of campaniform sensilla (CS) on the mesonotum, 2 (and sporadically 1) in *T. fuscipennis* and a lack of them in *T. sambuci*, is the characteristic that differentiates these species to the highest degree, though with several exceptions (0 CS in *T. fuscipennis* and 1 CS in *T. sambuci*). Additionally, the measured setae are shorter in specimens of the former species. Males of the latter are characterized by narrower area porosae on abdominal sternites V and VI and very often lack of them on sternite VII. Moreover, the analysis below declares eye length, distance between ocelli, and length of antennal segments

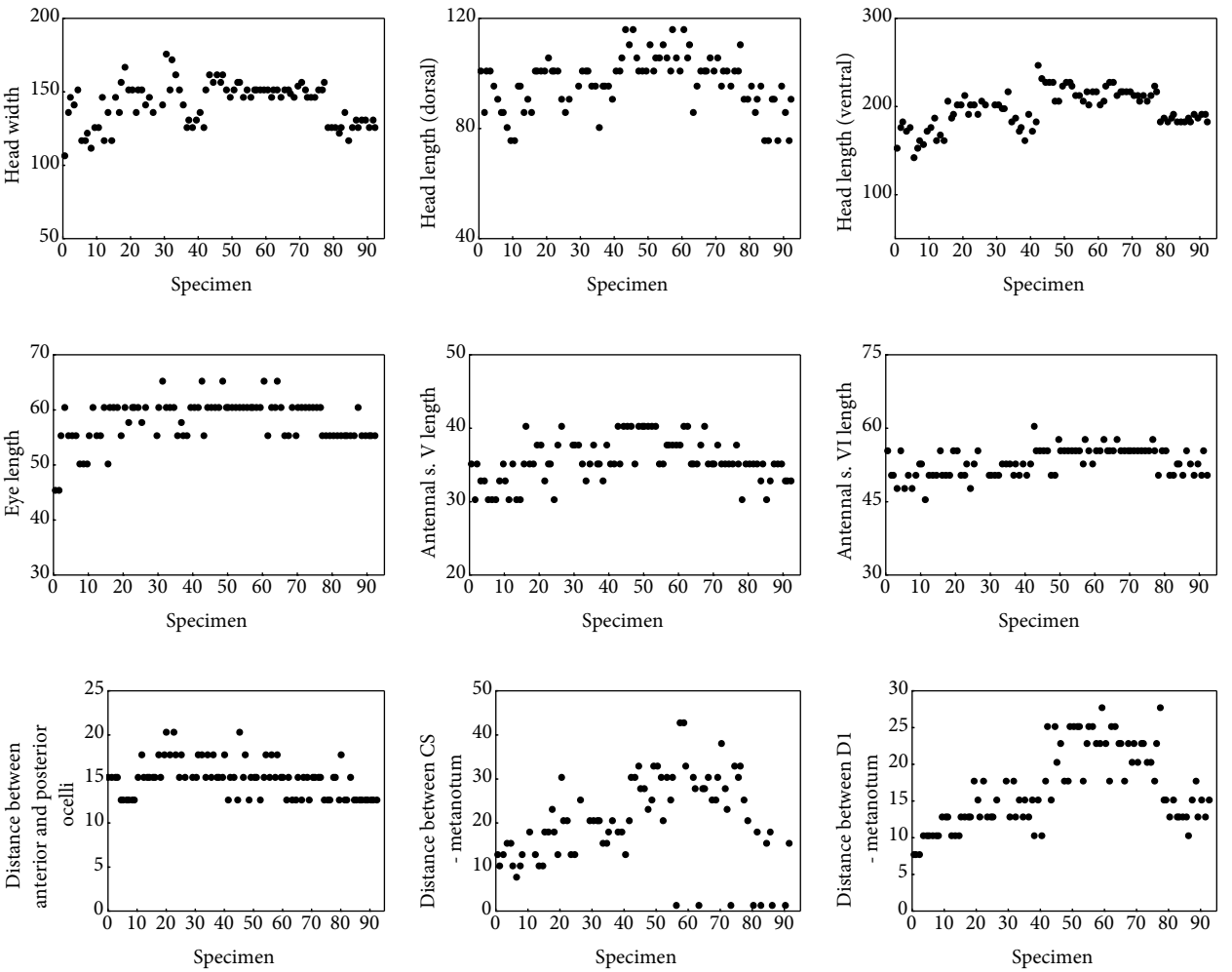

Figure 2a. Graphical visualization of characters 1–9. Axis x: 1–42 *Thrips fuscipennis*; 43–93 *T. sambuci* (length in μm).

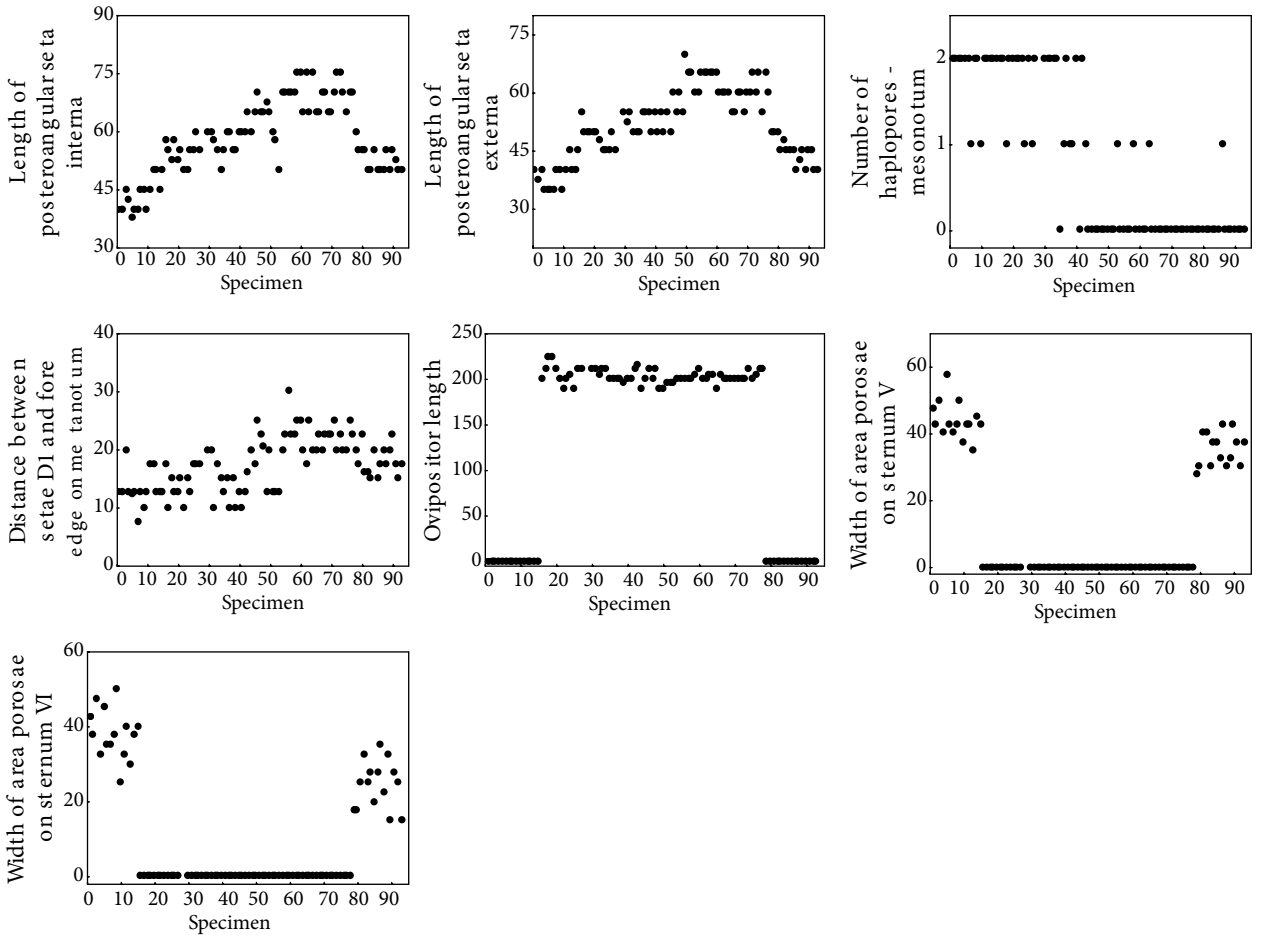

Figure 2b. Graphical visualization of characters 10–16. Axis x: 1–42 *Thrips fuscipennis*, 43–93 *T. sambuci* (length in µm).

as the characteristics with low significance in recognizing these species. Generally, no single character seems to be used for reliable identification.

Some of the characters (e.g., head width), despite their similar averages in both species (126.33 µm in *T. fuscipennis* male and 126.00 µm in *T. sambuci* male), significantly differ in their standard deviations (14.33 and 4.71), thus indicating a different degree of specific morphological plasticity. Generally, higher intraspecific variability refers to *T. fuscipennis*, for both females and males (Table 1), although some variables appear very similar in their standard deviations (e.g., head length and distance between anterior and posterior ocelli).

3.2. Correlation analysis

Correlation analysis, which refers to statistical relationships involving dependence, indicates a high correlation among the characters measured for *T. sambuci* (Table 2) and *T. fuscipennis* separately. The only exception corresponds with the distance between the posterior and anterior ocelli (insignificant correlation). The correlation analysis, calculated for both species simultaneously, has generalized the relationships among all the measured characters and emphasized a high degree of interactions between some of the characters, such as 15 and 16 (width of area porosae on sternum V and VI) or 10 and 11 (length of posteroangular seta interna and length of posteroangular seta externa), which undisputedly leads to a possible reduction in number of parallel variables. Final reduction in number of measured characters is required to speed up the identification process.

3.3. Factor analysis

The number of nonzero eigenvalues, estimating the number of linearly independent rows or columns, refers to the number of main sources of variability. This basic factor analysis (Table 3), which is to describe variability among observed factors, indicates that the first 2 eigenvalues are higher than 1 (those under 1 should usually be normally neglected), subsequently underlining the possible discriminability of both studied species, despite the similar values of their morphometric characters.

The value of rank equaling 2 (Figure 3) can be explained by the main variability among the characters

Table 2. Correlation among the measured characters. V - variable; 1 - head width; 2 - head length (dorsal side); 3 - head length (ventral side, including mouthcone); 4 - eye length; 5 - antennal segment V length; 6 - antennal segment VI length; 7 - distance between an anterior and posterior pair of ocelli; 8 - distance between CS and metanotum; 9 - distance between D1 and metanotum; 10 - length of posteroangular seta interna; 11 - length of posteroangular seta externa; 12 - number of CS on mesonotum; 13 - distance between setae D1 and fore edge on metanotum; 14 - ovipositor length; 15 - width of area porosae on sternum V; 16 - width of area porosae on sternum VI; and 17 - sex.

V	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
1	1.00	0.62	0.77	0.55	0.47	0.34	0.37	0.50	0.53	0.55	0.57	–0.16	0.47	0.72	–0.68	–0.66	0.71
2	0.62	1.00	0.53	0.48	0.47	0.41	0.38	0.52	0.53	0.52	0.56	–0.21	0.39	0.59	–0.55	–0.53	0.59
3	0.77	0.53	1.00	0.55	0.67	0.64	0.14	0.56	0.81	0.74	0.73	–0.45	0.67	0.69	–0.71	–0.71	0.70
4	0.55	0.48	0.55	1.00	0.44	0.18	0.27	0.36	0.52	0.55	0.52	–0.19	0.26	0.59	–0.61	–0.60	0.59
5	0.47	0.47	0.67	0.44	1.00	0.56	0.16	0.48	0.59	0.59	0.64	–0.25	0.37	0.63	–0.63	–0.63	0.63
6	0.34	0.41	0.64	0.18	0.56	1.00	–0.07	0.45	0.63	0.54	0.55	–0.45	0.53	0.41	–0.40	–0.43	0.41
7	0.37	0.38	0.14	0.27	0.16	–0.07	1.00	0.14	0.01	0.13	0.20	0.24	–0.04	0.40	–0.37	–0.33	0.39
8	0.50	0.52	0.56	0.36	0.48	0.45	0.14	1.00	0.52	0.54	0.58	–0.14	0.38	0.59	–0.57	–0.53	0.60
9	0.53	0.53	0.81	0.52	0.59	0.63	0.01	0.52	1.00	0.75	0.77	–0.55	0.69	0.59	–0.62	–0.62	0.60
10	0.55	0.52	0.74	0.55	0.59	0.54	0.13	0.54	0.75	1.00	0.88	–0.56	0.64	0.72	–0.75	–0.75	0.72
11	0.57	0.56	0.73	0.52	0.64	0.55	0.20	0.58	0.77	0.88	1.00	–0.44	0.63	0.76	–0.78	–0.77	0.77
12	–0.16	–0.21	–0.45	–0.19	–0.25	–0.45	0.24	–0.14	–0.55	–0.56	–0.44	1.00	–0.59	–0.10	0.19	0.26	–0.11
13	0.47	0.39	0.67	0.26	0.37	0.53	–0.04	0.38	0.69	0.64	0.63	–0.59	1.00	0.30	–0.32	–0.35	0.30
14	0.72	0.59	0.69	0.59	0.63	0.41	0.40	0.59	0.59	0.72	0.76	–0.10	0.30	1.00	–0.98	–0.94	1.00
15	–0.68	–0.55	–0.71	–0.61	–0.63	–0.40	–0.37	–0.57	–0.62	–0.75	–0.78	0.19	–0.32	–0.98	1.00	0.97	–0.98
16	–0.66	–0.53	–0.71	–0.60	–0.63	–0.43	–0.33	–0.53	–0.62	–0.75	–0.77	0.26	–0.35	–0.94	0.97	1.00	–0.94
17	0.71	0.59	0.70	0.59	0.63	0.41	0.39	0.60	0.60	0.72	0.77	–0.11	0.30	1.00	–0.98	–0.94	1.00

measured for both species, *T. fuscipennis* as well as *T. sambuci*. Surprisingly, despite high variability in the data set, the factor analysis clearly detects the main sources of variability evolving from the existence of 2 different species, as actually demonstrated in Figure 4, where the evaluation operates with females only to eliminate the role of sex in data distribution. Moreover, the same results can be obtained for males.

Table 3. Eigenvalue analysis of morphometric data for both species.

Eigenvalue (EV)	Total variance (%)	Cumulative EV	Cumulative variance (%)
9.036	56.477	9.036	56.477
1.960	12.247	10.996	68.725
0.887	5.543	11.883	74.267
0.796	4.977	12.679	79.244
0.616	3.848	13.295	83.092
0.560	3.502	13.855	86.594
0.524	3.276	14.379	89.870
0.417	2.606	14.796	92.476

3.4. Principal component and linear discriminant analyses

To follow the trends within the data matrix under the new dimension, principal component analysis (PCA), using the orthogonal transformation to convert observations into a set of principal components (PCs), was applied as projected between PC_1 vs. PC_2, PC_1 vs. PC_3, and PC_2 vs. PC_3 (Figure 5–7). Outcomes of this analysis prove the existence of 2

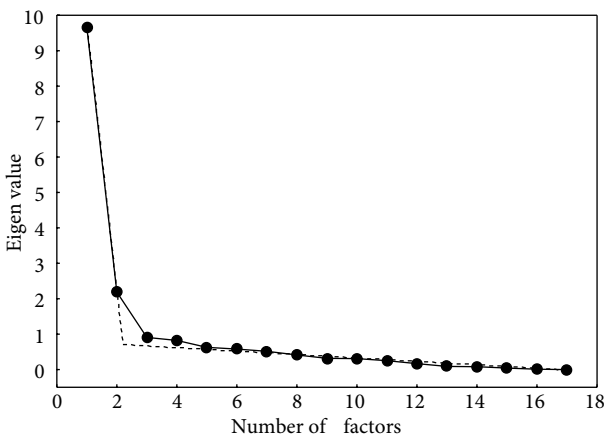

Figure 3. Plot of eigenvalues vs. number of eigenvalues factors.

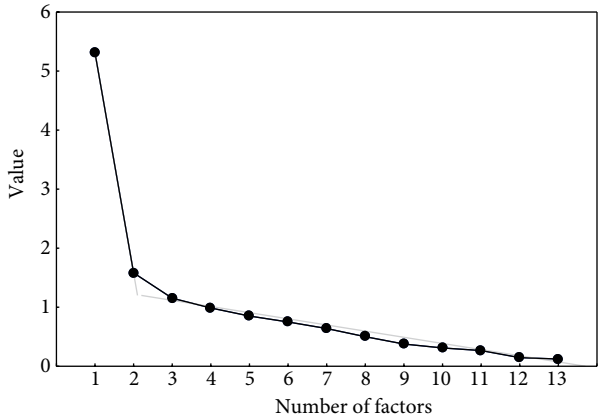

Figure 4. Eigenvalue analysis for *T. fuscipennis* and *T. sambuci* including females only; main variability is just given by 2 species, not by sex.

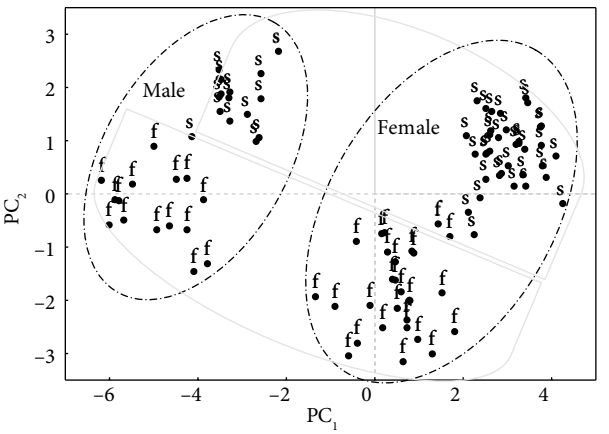

Figure 5. Graph of the principal components PC_1 vs. PC_2 concerning *Thrips* data analysis. f - *T. fuscipennis*; s - *T. sambuci*.

separate species, *T. fuscipennis* and *T. sambuci*, when all the specimens could be clearly distinguished, obviously with 2 distinct subgroups for each species, determined by sex. However, these subgroups are significantly less dispersed than the clusters concerning both species generally.

The same conclusion appears when applying linear discriminant analysis, which finds linear combinations of the variables measured. All the analyzed characters determine the species (Figure 8), with a strong impact of sex.

3.5. ANN analysis

Assuming a 3-layer ANN architecture and the RMS error function as determining the number of nodes in the hidden layer, the proposed model is visualized in Figure 9. The optimal architecture (17, *n*, 1) consists of 16 morphological characters plus sex as inputs, *n* number of nodes in the hidden layer, and 1 output depicted as 2 species (*T. sambuci* and *T. fuscipennis*). Consequently, 100% classification has been proven within the training process. Using cross-validation, the verification step appears successful when 1 to 10 specimens excluded from the database were tested. As a conclusion, the optimal ANN architecture, established as (17, 4, 1), is undisputedly able to distinguish both species reliably despite high variability of morphometric characters analyzed.

ANNs work in the system of hidden layers, which enables operations often unusual for standard identification tools. Thus, the relative values (ratio between various morphometric variables) appear more specific in distinguishing the species than their absolute values. When published online, the TRAJAN application software system enables semiautomated *T. sambuci* and *T. fuscipennis* discrimination. The user-friendly application analyzes all the input variables (morphometric characters) measured (in μm) as linear distances on digital images of slide-mounted specimens using any microscopic image analysis software. The reliability of the system has been successfully proven by a set of 52 independent specimens

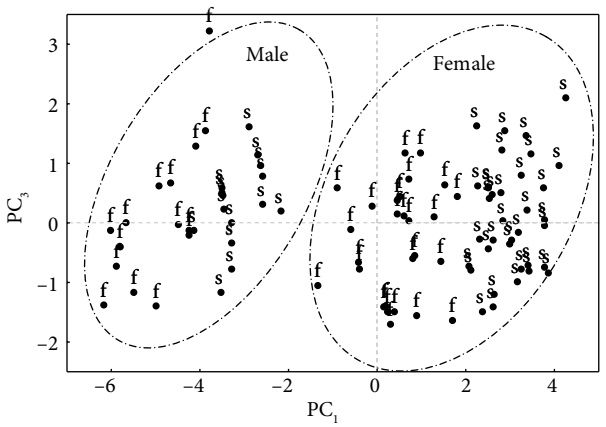

Figure 6. Graph of the principal components PC_1 vs. PC_3 concerning *Thrips* data analysis. f - *T. fuscipennis*; s - *T. sambuci*.

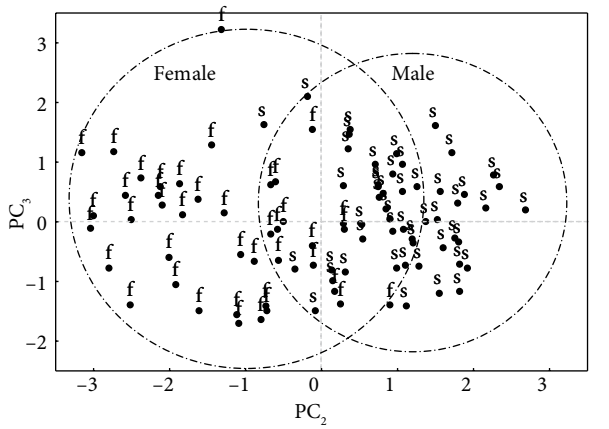

Figure 7. Graph of the principal components PC_2 vs. PC_3 concerning *Thrips* data analysis. f - *T. fuscipennis*; s - *T. sambuci*.

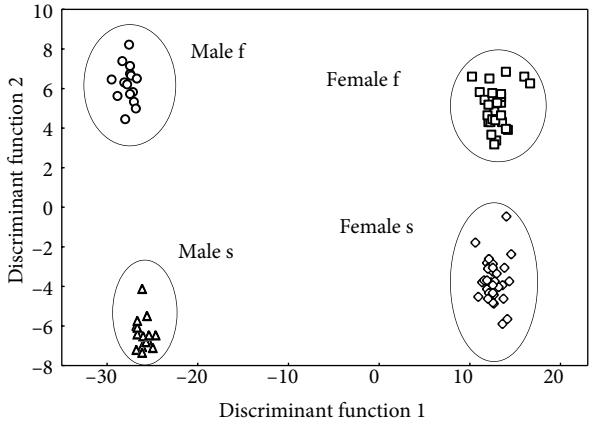

Figure 8. Linear discriminant analysis applied to distinguish *T. fuscipennis* – f and *T. sambuci* – s males and females.

with no errors recorded, even in the case of limited material. The application enables one to distinguish only the studied species, however; it indicates unknown results when the input data do not refer to either of the 2 defined alternatives. The first field expressing sex determines all other valid fields with relevant characters.

4. Discussion

Undisputedly, the use of artificially intelligent systems has spread to many fields of science (Weeks and Gaston, 1997; Do et al., 1999), including applied entomology (Solis-Sanchez et al., 2001; Fedor et al., 2009). ANNs, using a highly interconnected group of simulated neurons that process information in parallel (Haralabous and Georgakarakos, 1996) and learning from a set of examples, have been widely theoretically described (Freeman and Skapura, 1992; Ripley, 1993; Haykin 1994; Haralabous and Georgakarakos, 1996); therefore, there is no need to do so in this methodical paper.

Being quite different from standard statistical tools, the approach is unique and autonomous; however, it is often described as a black-box system with no readily interpretable explanation for the prediction provided (Ripley, 1993; Haralabous and Georgakarakos, 1996). Despite entomologists often trying to find a parallel in basic statistics (average, maximum, minimum, median), the process for obtaining the internal structure of a network is complex, with no defined methodology (Isasi and Galván, 2004). As an advantage, the quantity of the material reflects the specific approach of artificially intelligent systems, when even a limited number of samples appear sufficient for reliable analysis and searching for patterns. This differs from the standard statistical tools, operating with thousands of samples, which have, in fact, practical consequences for the development of identification systems. In Thysanoptera identification, this phenomenon has been proven by Fedor et al. (2008, 2009).

This paper has shown the practical use of artificial intelligence in applied pest identification systems, having been encouraged by many previous papers dealing with various species discrimination (Clark, 2003; Kavdır, 2004; Marini et al., 2004; Esteban et al., 2009). We have defined a total of 17 characters, measured or determined for each *Thrips fuscipennis* and *T. sambuci* specimen, including 16 quantitative morphometric or qualitative characters on different parts of their bodies and sex as the 17th. Most of the characters have been commonly applied in *Thrips* species identification (e.g., Nakahara, 1994; zur Strassen, 2003; Mound and Masumoto, 2005; Mound and Azidah, 2009) and have been supposed to enable distinguishing between Thysanoptera species (Fedor et al., 2008, 2009). Generally, there are many characters available for analyzing among the *Thrips* species, such as those published by Kucharczyk and Kucharczyk (2009) in their taxonomic revision of *T. atratus* and *T. montanus* (e.g., the number of distal setae on the first vein of the forewing, the shape of the microtrichial comb on the posterior margin of tergum VIII, the number of discal setae on abdominal sternites V and VII, and the length of antennal stylus). The high

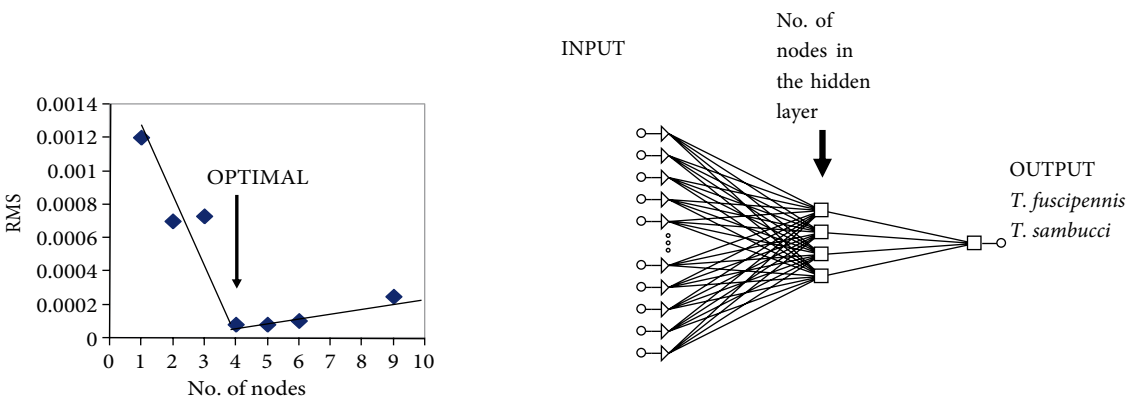

Figure 9. RMS as a function of the number of nodes in the hidden layer (left) and suggested optimal architecture (17, 4, 1).

value of the selected and measured characters (especially morphometric) has been recently proven in our previous projects on some other Thysanoptera taxa (Fedor et al., 2008, 2009), including those controlled and monitored by phytosanitary staff. Mature *Thrips sambuci* and *T. fuscipennis* may be distinguished by differences in color of antennal segments (Schliephake and Klimt, 1979; zur Strassen, 2003). However, according to our experiences, the color characters can be very variable in specimens originating from different populations and stations, and sometimes it is not possible to accurately identify the species (Mound and Minaei, 2010). Moreover, for ANN systems to establish a color scale with clear borders may be quite challenging.

Our project has been predominantly established on a set of morphometric characters, which used to be presented just as the additional data in many identification keys. This paper, moreover, has emphasized morphometric characters in their appropriate combination as an autonomous source of information necessary for reliable discrimination. Despite intraspecific variability induced both genetically and ecologically, when appropriately combined (not single), morphometric characters are capable of being successfully applied in identification systems (e.g., Vaňhara et al., 2007; Esteban et al., 2009; Muráriková et al., 2011).

The use of an optimal 3-layer ANN architecture (17, 4, 1) enables fast and reliable classification, with nearly 100% accuracy shown during the extensive verification process. Compared to many other similar systems published by Marini et al. (2004), Kavdır (2004), Vaňhara et al. (2007), Fedor et al. (2008), and Han et al. (2012), the proposed optimum architecture works very reliably. For instance, to achieve the differentiation of 2 *Juniperus* species, a feedforward MLP network was proposed, which attained 98.6% success in the training group and 92.0% success in the testing or unknown group (Esteban et al., 2009).

The data set generated during the morphometric measurements represents a valuable source for species discrimination. Although this paper does not offer any taxonomic revision, the key to reliable ANN analysis lies in an appropriate matrix transformation and selection of specific statistical approach, which applies in parallel several autonomous methods to prove the hypotheses more reliably. For instance, PCA is one of the simplest eigenvector-based multivariate analyses, and its operation can be thought of as revealing the internal structure of the data in a way that best explains the variance in the data (Chiapella, 2000; Lilburn and Garrity, 2004; Apuan et al., 2010). The method was thus successfully used to determine the relations among 26 morphological characters (e.g., length of antennal segments and sense cones, length of dorsal setae on head, length of pronotal setae) in a group of 35 second larval instar *Thrips* species from Central Europe (Kucharczyk, 2010). The PCA method was applied to distinguish similar *Thrips* (*T. atratus* and *T. montanus*) species for the first time by Kucharczyk and Kucharczyk (2009). Since the description, both of them were reclassified to *Thrips*, *Taeniothrips*, or *Similothrips* genera (Priesner, 1964; Schliephake and Klimt, 1979; Schliephake, 2001). Finally, 8 female and 12 male morphological characters were successfully analyzed to recognize these species more easily. The multidimensional methods (such as PCA or factor analysis) were applied in this paper as an important step in ANN processing, to prove whether the set of selected variables had the power to reliably distinguish both species in artificially intelligent systems, as preliminarily proposed in our previous papers (Fedor et al., 2008, 2009), including in the case of *T. sambuci* and *T. fuscipennis* PCA discrimination (Kucharczyk et al., 2012).

In the factor analysis, the low cumulative variance (74.27%) for the first 3 eigenvalues emphasizes some other sources of variability or noise in the data. The reason may be a wider spread of morphometric character values caused by high morphological plasticity and/or simply errors in the measurements. Moreover, high noise in the variables measured can correspond with some characters with no significant contribution in distinguishing the species.

The ANN system may be successfully applied as a credible supplementary (alternative) identification tool, for instance, when the standard characters have high overlap. If the common morphological keys hint at practical problems of high morphological plasticity, the artificially intelligent system is capable of extracting hidden information from highly noisy data. One of the main objectives of this paper was to offer ANNs as a reliable identification system for practical, specific phytosanitary use, when the other standard methods appear limited financially (e.g., DNA analysis) or by high morphological plasticity (common morphological keys). Artificially intelligent systems could refine their performance by comparing multivariate continuous morphometric data more efficiently or simply providing an independent check, e.g., for specific cases and critical taxa (Clark, 2003; Esteban et al., 2009), specimens partly damaged by rough collecting methods (sticky traps, aeroplanktonic traps, Tullgren photoeclectors, etc.), or old slides (Fedor et al., 2008).

Homologous characters can be objectively defined as distances on the thrips body, and the measurement of such distances requires only limited experience in slide-mounting techniques, as well as basic knowledge of thrips morphology. Technical requirements are limited to a microscope with an eye-piece graticule or a digital camera and image analysis software, which would allow more comfort.

Identification of insect species can be sometimes time-consuming and can require technical expertise, so an automated insect identification method is needed

(Han et al., 2012). Artificial intelligence offers reliable and independent systems, but they require a sufficient software background and at least elementary statistical experience. Therefore, any applied research should be accompanied by the effective software proposal as a final product of ANN application in species identification. Although the need for the automation of routine biological object identification has been rather concentrated into image analysis software tools (Weeks and Gaston, 1997; Do et al., 1999; Mayo and Watson, 2007) or even the classification system design based on Blackfin DSP and 3G wireless communication technology, which is composed of a remote online classification platform with a digital signal processor and a host control platform (Han et al., 2012), the system proposed in this paper, which is based on morphometric variables, works with greater reliability.

Obviously, ANN systems, as any alternative approach to species identification, have their specific limits (Weeks and Gaston, 1997; Gaston and O'Neill, 2004; Fedor et al., 2008). Selection of significant morphometric characters markedly depends on the analyzed taxa and should be thus performed by a professional taxonomist (theoretical phase). The power of artificially intelligent tools may rely on each member of the phytosanitary staff and his preparatory and measurement skills (practical phase). Obviously it takes some time to get basic experience when using ANN systems. For instance, errors in identification of Thysanoptera can be caused by intraspecific variation, even though distinguishing between 2 disputably different species occurs almost with no problems. The ANN architecture in this case enables one to distinguish between *T. sambuci* and *T. fuscipennis* only. Specimens of other species will not be accepted in the identification window.

The artificially intelligent systems may be practically applied as a credible alternative for online semiautomated pest identification. The online user-friendly application clearly enables prompt *T. sambuci* and *T. fuscipennis* discrimination according to the set of the analyzed (mainly morphometric) characters; its high reliability has been successfully demonstrated by a set of 52 independent specimens (measured by independent persons) with no errors recorded. The system has been established to distinguish 2 analyzed species, but it also possesses an ability to indicate when a species is unknown or when it encounters invalid data. Conceptually, this approach should be applied as a supplementary method in specific cases for a group of several relatively disputably discriminable species. Obviously, we do not argue that the ANN identification systems could automatically replace all the standard methods available in applied entomology; rather, our methodical paper presents a tool that may find practical use in some specific problems of reliable and prompt identification of pests.

Acknowledgements

Financial support for this project was provided by the Slovak Science Foundation VEGA 1/0137/11, ITMS 26240220086 and by European Social Fund CETPO project – CZ 1.07/2.3.00/20.0166.

References

Aldrich BT, Magirang EB, Dowell FE, Kambhampati S (2007). Identification of termite species and subspecies of the genus *Zootermopsis* using near-infrared reflectance spectroscopy. J Insect Sci 7: 18.

Ananthakrishnan TN (2005). Perspectives and dimensions of phenotypic plasticity in insects. In: Ananthakrishnan TN, Whitman D, editors. Insect Phenotypic Plasticity: Diversity of Responses. Enfield, NH, USA: Science Publishers, pp. 1–23.

Apuan DA, Torres MAJ, Demayo CG (2010). Describing variations and taxonomic status of earthworms collected from selected areas in Misamis Oriental, Philippines using principal component and parsimony analysis. Egypt Acad J Biolog Sci B Zoology 2: 27–36.

Bilgili M (2011). The use of artificial neural networks for forecasting the monthly mean soil temperatures in Adana, Turkey. Turk J Agric For 35: 83–93.

Brunner PC, Fleming C, Frey JE (2002). A molecular identification key for economically important thrips species (Thysanoptera: Thripidae) using direct sequencing and a PCR-RFLP-based approach. Agr Forest Entomol 4: 127–136.

Chiapella J (2000). The *Deschampsia cespitosa* complex in central and northern Europe: morphological analysis. Bot J Linn Soc 134: 495–512.

Clark JY (2003). Artificial neural networks for species identification by taxonomists. Biosystems 72: 131–147.

Do MT, Harp JM, Norris KC (1999). A test of a pattern recognition system for identification of spiders. B Entomol Res 89: 217–224.

Esteban LG, Fernández FG, de Palacios P, Romero RM, Cano NN (2009). Artificial neural networks in wood identification: the case of two *Juniperus* species from the Canary islands. IAWA J 30: 87–94.

Fausett L (1994). Fundamentals of Neural Networks: Architectures, Algorithms and Applications. New York, NY, USA: Prentice Hall.

Fedor P, Malenovský I, Vaňhara J, Sierka W, Havel J (2008). Thrips (Thysanoptera) identification using artificial neural networks. B Entomol Res 98: 437–447.

Fedor P, Vaňhara J, Havel J, Malenovský I, Spellerberg I (2009). Artificial intelligence in pest insect monitoring. Syst Entomol 34: 398–400.

Frantz G, Mellinger HC (1997). THRIPS, a computerized knowledgebase for the identification and management of Thrips infesting vegetables in the United States. Proc Fla State Hort Soc 128: 1–4.

Freeman JA, Skapura DM (1992). Neural Network: Algorithm, Applications, and Programming Techniques. Reading, MA, USA: Addison-Wesley.

Gaston KJ, O'Neill MA (2004). Automated species identification: why not? Philos T Roy Soc B 359: 655–667.

Han R, He Y, Liu F (2012). Feasibility study on a portable field pest classification system design based on DSP and 3G wireless communication technology. Sensors 12: 3118–3130.

Haralabous J, Georgakarakos S (1996). Artificial neural networks as a tool for species identification of fish schools. ICES J Mar Sci 53: 173–180.

Haykin S (1994). Neural Networks - A Comprehensive Foundation. New York, NY, USA: Macmillan College Publishing Company.

Henneberry TJ, Taylor EA, Smith FF (1961). Foliage and soil treatments for control of flower thrips in outdoor roses. J Econ Entomol 54: 233–235.

Isasi P, Galván IM (2004). Redes Neuronales Artificiales, un Enfoque Práctico. Madrid, Spain: Pearson Educación S.A. (in Spanish).

Jacobson RJ (1995). IPM in cucumbers – prevention of establishment of *Frankliniella occidentalis*. Mededelingen van de Faculteit van de Landbouwwntenschappen 60: 857–863.

Kavdır İ (2004). Discrimination of sunflower, weed and soil by artificial neural networks. Comput Electron Agr 44: 153–160.

Kucharczyk H (2010). Comparative Morphology of the Second Larval Instar of the *Thrips* Genus Species (Thysanoptera: Thripidae) Occurring in Poland. Olsztyn, Poland: Mantis.

Kucharczyk H, Kucharczyk M (2009). *Thrips atratus* Haliday, 1836 and *Thrips montanus* Priesner, 1920 (Thysanoptera: Thripidae) – one or two species? Comparative morphological studies. Acta Zool Acad Sci H 55: 349–364.

Kucharczyk H, Kucharczyk M, Stanislawek K, Fedor P (2012). Application of PCA in taxonomy research – Thrips (Insecta, Thysanoptera) as a model group. In: Sanguansat P, editor. Principal Component Analysis - Multidisciplinary Applications. New York, NY, USA: InTech, pp. 111–126.

Lilburn TG Garrity GM (2004). Exploring prokaryotic taxonomy. Int J Sys Evol Micr 54: 7–13.

Marini F, Zupan J, Magri A (2004). On the use of counterpropagation artificial neural networks to characterize Italian rice varieties. Anal Chim Acta 510: 231–240.

Mayo M, Watson AT (2007). Automatic species identification of live moths. Knowl-Based Sys 20: 195–202.

Mehle N, Trdan S (2012). Traditional and modern methods for the identification of thrips (Thysanoptera) species. J Pest Sci 85: 179–190.

Moore A, Miller RH (2002). Automated identification of optically sensed aphid (Homoptera: Aphidae) wingbeat waveforms. Ann Entomol Soc Am 95: 1–8.

Moritz G, Delker C, Paulsen M, Mound LA, Burgermeister W (2000). Modern methods for identification of Thysanoptera. EPPO Bulletin 30: 591–593.

Moritz G, Morris DC, Mound LA (2001). ThripsID – Pest Thrips of the World. CD ROM. Melbourne, Australia: ACIAR, CSIRO Publishing.

Moritz G, Mound LA, Morris DC, Goldarazena A (2004). Pest Thrips of the World, Visual and Molecular Identification of Pest Thrips. CD ROM. Brisbane, Australia: Centre for Biological Information Technology.

Mound LA (2005). Fighting, flight and fecundity: behavioural determinants of Thysanoptera structural diversity. In: Ananthakrishnan TN, Whitman D, editors. Insect Phenotypic Plasticity: Diversity of Responses. Enfield, NH, USA: Science Publishers, pp. 81–105.

Mound LA (2010). Species of the genus *Thrips* (Thysanoptera, Thripidae) from the Afro-tropical Region. Zootaxa 2423: 1–24.

Mound LA, Azidah AA (2009). Species of the genus *Thrips* (Thysanoptera) from Peninsular Malaysia, with a checklist of recorded Thripidae. Zootaxa 2023: 55–68.

Mound LA, Kibby G (1998). Thysanoptera: An Identification Guide. Wallingford, UK: CAB International.

Mound LA, Masumoto M (2005). The genus *Thrips* (Thysanoptera, Thripidae) in Australia, New Caledonia and New Zealand. Zootaxa 1020: 1–64.

Mound LA, Minaei K (2010). Taxonomic problems in character state interpretation: variation in the wheat thrips *Haplothrips tritici* (Kurdjumov) (Thysanoptera, Phlaeothripidae) in Iran. Deut Entomol Z 57: 233–241.

Muráriková N, Vaňhara J, Tóthová A, Havel J (2011). Polyphasic approach applying artificial neural networks, molecular analysis and postabdomen morphology to West Palaearctic *Tachina* spp. (Diptera, Tachinidae). B Entomol Res 101: 165–175.

Nakahara S (1994). The Genus *Thrips* Linnaeus (Thysanoptera: Thripidae) of the New World. Technical Bulletin 1822. Washington, DC, USA: United States Department of Agriculture.

Patterson D (1996). Artificial Neural Networks: Theory and Applications. Singapore: Prentice Hall.

Priesner H (1964). Ordnung Thysanoptera (Fransenflügler, Thripse). Berlin, Germany: Akademie-Verlag (in German).

Ripley BD (1993). Statistical aspects of neural networks. In: Barndoff-Nielsen OE, Cox DR, Jensen JL, Kendall WS, editors. Chaos and Networks – Statistical and Probabilistic Aspects. London, UK: Chapman and Hall, pp. 40–123.

Rugman-Jones PF, Hoddle MS, Mound LA, Stouthamer R (2006). Molecular identification key for pest species of *Scirtothrips* (Thysanoptera: Thripidae). J Econ Entomol 99: 1813–1819.

Schliephake G (2001). Verzeichnis der Thysanoptera (Fransenflügler) – Physopoda (Blasenfüße) – Thrips Deutschlands. Entomof Germ 5: 91–106 (in German).

Schliephake G, Klimt K (1979). Thysanoptera. Die Tierwelt Deutschlands 66. Jena, Germany: G. Fisher Verlag (in German).

Schuetz I, Breitenstein A, Moritz G (2010). Computer-based identification key for pest thrips and tospoviruses by use of LucID 3.4, ITS-RFLP and low-density BioChip technology. In: Persley D, Wilson C, Thomas J, Sharman M, Tree D, editors. Proceedings of the IXth International Symposium on Thysanoptera and Tospoviruses, 31 August – 4 September, 2009. J Insect Sci 10: 1–58.

Solis-Sanchez LO, Castañeda-Miranda R, García-Escalante JJ, Torres-Pacheco I, Guevara-González RG, Castañeda-Miranda CL, Alaniz-Lumbreras PD (2001). Scale invariant feature approach for insect monitoring. Comp Electron Agr 75: 92–99.

Taylor M (2010). Latest lucid identification tools. In: Persley D, Wilson C, Thomas J, Sharman M, Tree D, editors. Proceedings of the IXth International Symposium on Thysanoptera and Tospoviruses, 31 August – 4 September, 2009. J Insect Sci 10: 1–58.

Toda S, Komazaki S (2002). Identification of thrips species (Thysanoptera: Thripidae) on Japanese fruit trees by polymerase chain reaction and restriction fragment length polymorphism of the ribosomal ITS2 region. B Entomol Res 92: 359–363.

Tohidi M, Sadeghi M, Mousavi SR, Mireei SA (2012). Artificial neural network modeling of process and product indices in deep bed drying of rough rice. Turk J Agric For 36: 738–748.

Vaňhara J, Murárikova N, Malenovský I, Havel J (2007). Artificial neural networks for fly identification: a case study from the genera *Tachina* and *Ectophasia* (Diptera, Tachinidae). Biologia 62: 462–469.

Vierbergen G, Kucharczyk H, Kirk WDJ (2010). A key to the second instar larvae of the Thripidae of the Western Palaearctic region (Thysanoptera). Tijdschr Entomol 153: 99–160.

Weeks PJD, Gaston KJ (1997). Image analysis, neural networks, and the taxonomic impediment to biodiversity studies. Biodivers Conserv 6: 263–274.

Zhang WJ, Barrion AT (2006). Function approximation and documentation of sampling data using artificial neural networks. Environ Monit Assessment 122: 185–201.

zur Strassen R (2003). Die terebranten Thysanopteren Europas und des Mittelmeer-Gebietes. Die Tierwelt Deutschlands 74: 1–271 (in German).