Systematic Entomology

Systematic Entomology (2009), DOI: 10.1111/j.1365-3113.2008.00461.x

METHODS

Artificial intelligence in pest insect monitoring

PETER FEDOR[1], JAROMÍR VAŇHARA[2], JOSEF HAVEL[2], IGOR MALENOVSKÝ[3] and IAN SPELLERBERG[4]

[1]Department of Ecosozology, Faculty of Natural Sciences, Comenius University, Bratislava, Slovak Republic, [2]Department of Botany and Zoology and Department of Chemistry, Faculty of Science, Masaryk University, Brno, Czech Republic, [3]Department of Entomology, Moravian Museum, Brno, Czech Republic and [4]Isaac Centre for Nature Conservation, Lincoln University, Canterbury, New Zealand

Abstract. Global problems of hunger and malnutrition induced us to introduce a new tool for semi-automated pest insect identification and monitoring: an artificial neural network system. Multilayer perceptrons, an artificial intelligence method, seem to be efficient for this purpose. We evaluated 101 European economically important thrips (Thysanoptera) species: extrapolation of the verification test data indicated 95% reliability at least for some taxa analysed. Mainly quantitative morphometric characters, such as head, clavus, wing, ovipositor length and width, formed the input variable computation set in a Trajan neural network simulator. The technique may be combined with digital image analysis.

Introduction

In response to a worsening in the global food situation, scientists have been seeking innovative ways to control the impact of agriculture pests by developing integrated production systems involving biotechnology, including genetically modified organisms and new generation pesticides. We tested the reliability of a new method for prompt and effective identification and monitoring of pest species through analysis of mainly morphometric characters and interspecific variation.

Artificial neural networks (ANN), defined under artificial intelligence, suggested possible practical methods for semi-automated identification of biological objects (MacLeod, 2008). Together with other statistical tools, such as principal component analysis or classification trees, such networks meet digital age science. An optimum architecture, established and supervised by experts, transforms metadata through multilayer system processing, artificially intelligent apparatus.

Economically important European thrips (Thysanoptera) of 101 species (Supporting Information ST1), including European and Mediterranean Plant Protection Organization quarantine-listed and tospovirus-transmitting pests, such as *Thrips palmi* or *Frankliniella occidentalis*, were evaluated to develop a precise and reliable digital instrument for identification. Experiments with several ANN types (radial basis function, linear, probabilistic and multilayer perceptrons networks) suggested multilayer perceptrons as efficient for such a purpose generally. Its architecture was constructed conventionally with three or more feed-forward layers, i.e. input, output and one or several hidden layers.

More than 3000 specimens, both males and females, were analysed in accordance with standard preparatory techniques. For instance, 17 quantitative morphometric characters (measured as linear distances on digital images), such as head, clavus, wing, ovipositor length and width (Fig. 1), two qualitative two-state characters (presence/absence) and sex (Fedor *et al.*, 2008) formed the input variable computation set in a Trajan neural network simulator (Trajan Software, Ltd, 1996–1998) for cereal damaging and grain damage causing thrips of the *Limothrips* genus. The experimental uncertainty for the digital images was 0.03 μm. In the training process (equivalent to fitting the model) with a tendency to minimize the root mean square value, we sought an optimal architecture with the appropriate number of layers and nodes. A verification set of randomly selected specimens proved a strong generalization in processing new samples.

Reliable species distinction would not be possible if a single character only was considered. However, in appropriate combinations (relative values), character states can be unique

Correspondence: Peter Fedor, Faculty of Natural Sciences, Comenius University, Mlynska dolina, 842 15 Bratislava, Slovak Republic. E-mail: fedor@fns.uniba.sk

Unpublished for the purposes of zoological nomenclature (Art. 8.2, ICZN)

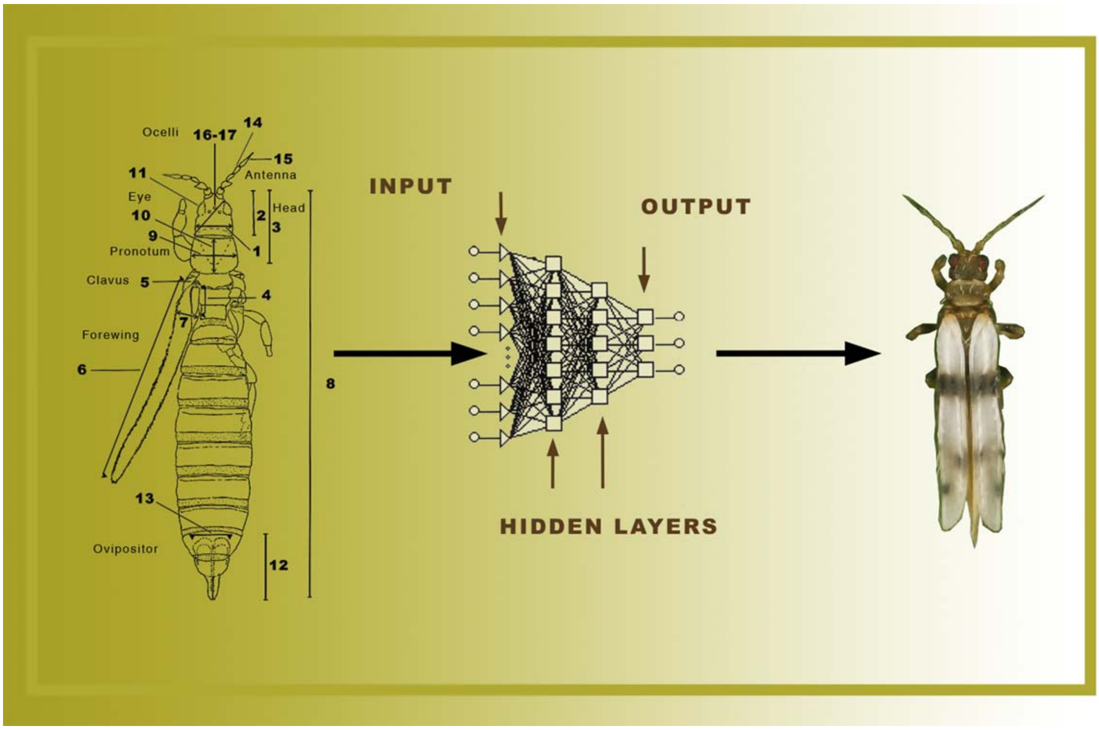

Fig. 1. An operating artificial neural network system in pest identification. Morphometric and qualitative variables (1, head width; 2, head length – dorsal side; 3, head length – ventral side; 4, clavus length; 5, clavus width; 6, forewing length; 7, forewing basal width; 8, total body length; 9, pronotum width; 10, pronotum length; 11, eye length; 12, ovipositor length; 13, ovipositor width; 14, antennal segment V length; 15, antennal segment VI length; 16, distance between the posterior pair of ocelli; 17, distance between an anterior and posterior pair of ocelli; 18, forewing with three light bands –yes/no; 19, clavus distinctly widest basally – yes/no) are used as input data. In a true artificial neural network architecture, input data are analysed through several hidden layers to yield the final identification. The ANN model including its architecture and photograph is a demonstrative example.

and specific. Detailed ANN analyses on 18 common European Thysanoptera species have been published recently (Fedor *et al.*, 2008). An extrapolation of the data obtained from the verification tests indicated that for a 95% correct identification, a training set with 300–350 specimens in total and at least six specimens of each species is sufficient.

In such a large database, lower subsystems can be established and analysed separately (e.g. Panchaetothripinae, *Frankliniella* spp., *Thrips* spp.), while also shaping a complex supersystem with high identification reliability. In combination with semi-automated digital image-encoding software and determination according to the most reliable characters, the method may find a wide practical application in phytosanitary work.

Sampling and establishing the appropriate database appears more laborious than final analyses and syntheses, but our experiences indicate the possibility of using just a few specimens if necessary. A wide character range even enables the identification of damaged insects where traditional dichotomous keys are useless. Optimally, ANNs should be combined with various multi-access keys (e.g. Moritz *et al.*, 2001).

The practical use of ANNs for pest identification has several limitations (Gaston & O'Neill, 2004). For thrips, problems may originate in intraspecific variation (e.g. wing and sex polymorphism). However, enough evidence exists that precise and effective pest monitoring could have far-reaching implications for farmers and food policy without harming biodiversity, expensive inputs, or plaguing human dignity. The use of ANNs corresponds with more effective control of harmful pests and thus with prevention of serious crop damage.

Supporting Information

Additional Supporting Information may be found in the online version of this article under the DOI reference: doi: 10.1111/j.1365-3113.2008.00461.x.

ST1 International Society for Pest Information and European and Mediterranean Plant Protection Organization thrips pests list for Europe measured for the artificial neural network analysis.

Please note: Wiley-Blackwell are not responsible for the content or functionality of any supporting information supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.

Acknowledgements

We acknowledge VEGA 1/4339/07, MSM 0021622416, GAČR 525/06/0663, MK 00009486201 Dr Laurence Mound for his kind comments and Mr M. Deml for his photographic skills.

References

Fedor, P., Malenovský, I., Vaňhara, J., Sierka, W. & Havel, J. (2008) Thrips (Thysanoptera) identification using artificial neural networks. *Bulletin of Entomological Research*, **98**, 437–447.

Gaston, K.J. & O'Neill, M.A. (2004) Automated species identification: why not? *Philosophical Transactions of the Royal Society of London, Series B*, **359**, 655–667.

MacLeod, N. (ed.) (2008) *Automated Taxon Identification in Systematics. Theory, Approaches and Applications*. CRC Press, Boca Raton.

Moritz, G., Morris, D.C. & Mound, L.A. (2001) *Thrips ID – Pest Thrips of the World*. CD ROM. ACIAR, CSIRO Publishing, Collingwood, Australia.

Trajan Software, Ltd (1996–1998) *Trajan Neural Network Simulator, Version 3.0D, Computer Program*. Trajan Software, Horncastle, Lincolnshire.

Accepted 3 December 2008