JOURNAL OF PLANT PROTECTION RESEARCH
Vol. 41, No. 4 2001

CHEMOMETRICS AS A TOOL FOR PLANT – INSECT ECOLOGICAL STUDIES

JANUSZ SZAFRANEK[1], JAN NAWROT[2], BEATA SZAFRANEK[1], EDMUND MALIŃSKI[1]

[1]Department of Environmental Analysis, University of Gdańsk,
Sobieskiego 18, 80-952 Gdańsk, Poland
[2]Institute of Plant Protection,
Miczurina 20, 60-318 Poznań, Poland
e-mail:janat@chemic.chem.univ.gda.pl

Accepted: May 15, 2001

Abstract: The process of plant selection by insects is mediated by repellents and attractants. Several compounds may be involved in this interaction. Thus intraspecific variation of the compounds concentration play an important role in the herbivory. The best tool for the characteristic of this variation is chemometrics. The strategy of the analysis with the use of literature data on terpenes and sesquiterpenes variations in *Pinus caribaea* needles in relation to *Atta laevigata* herbivory is exemplified herein. Simple cluster analysis and principal components analysis were used for the data study. Two factors were found to be sufficient to describe total variation in more than 90%. Factor 1 is responsible for repellent properties. From factor loading, the relevant chemical compounds were identified.

Key words: chemometrics, terpenes, sesquiterpenes, pine needles, ants

I. INTRODUCTION

Many plant chemicals of different structures control the insect behaviour. Insect chemoreceptors recognise them as attractants or repellents. Ecological interaction is a complex biological relationship being modified by synergism for compounds in the mixtures. For understanding of these ecological interactions, qualitative and quantitative full description of a system (composition) is thus needed. This is also true for intra-species interaction controlled by pheromones. Specific mixture composition is usually responsible for the bioactivity. Thus the full strategy of bioactivity consists of the following stages shown below.

Strategy {
Biological phenomen
⇓
Chemical analysis
⇓
Chemometrics
⇓
Conclusion

There are many chemometrics methods but not all are able to process all kind of analytical data. Therefore the type of variables (continuous, discrete, order) already limits the choice of the methods.

The main aim of this presentation is to show how one can obtain the conclusions in chemical studies of epicuticular waxes by relevant chemometrics methods. We use herein literature analytical data of mono- and sesquiterpene compositions of defoliated and undefoliated trees *Pinus caribaea* by ants *Atta laevigata* (Barnola 1994). The data have been originally evaluated only by discrimination analysis. Absolute frequency distribution of the canonical variable in the discriminant analysis distinguished defoliated and undefoliated pine groups of the same species. January analysis of newly defoliated pines revealed that the concentration of myrcene and caryophyllene were lower than in undefoliated trees suggesting terpenes involvement in plant – insect interaction (Howard et al. 1988). Thus interspecific terpene variation in pine needles should be carefully studied.

II. MATERIALS AND METHODS

Chemical analysis data of epicuticular waxes of defoliated and undefoliated trees were taken from paper of Barnola et al. (1994). The quantitative terpene and sesquiterpene compositional results were analysed by chemometrics' methods. Data were transformed and autoscaled (Kowalski and Bender 1972) before analysis. Cluster analysis and principal component routines were used. EXCEL, STATISTICA and MATLAB softwares were applied.

According to chemical analysis results, sixteen properties (variables) were assigned for all cases (trees) which gave data vectors. Cluster analysis classified the objects (cases) on their similarities (Massart and Kaufman 1983) in p-space. Principal component analysis was used to reduce the dimensions of the space (Mardia et al. 1979).

III. RESULTS AND DISCUSSION

Sixteen terpene analysis results of defoliated and undefoliated trees are shown in table 1. A search for natural grouping of the samples is a primary stage of any data analysis. Cluster analysis was applied with data set of Euclidean distance squares as a similarity matrix. Figure 1 shows the cluster of the trees. Both sets, defoliated and undefoliated trees, are distinguished with high degree of similarities in the groups. Higher variances are, however, observed in the group of undefoliated trees. But no conclusion could be drawn about chemicals involved in the interaction. Herein, the correlation techniques of n-variables in n-dimensional space are needed.

Factor analysis (FA) is a solution of the problem because it describes the correlation between variables in a data set by means of small number of factors. The factors, which describe 90% of the total variances, are sufficient for a proper correlation. Table 2 contains eigenvalues and variances related to factors of the data studied.

Table 1

Terpenes and sesquiterpenes in pine needles defoliated and undefoliated [μg/g]. Data from Barnola et al. 1994.

Trees	1	2	3	4	5	6	7	8
defoliated-J	115.2	5.1	133.9	51.1	164.4	460.7	7.2	6.0
undefoliated-J	113.0	5.7	64.3	76.4	222.1	631.1	7.4	3.7
defoliated-M	155.1	5.6	153.1	53.3	146.1	475.0	12.5	7.0
undefoliated-M	172.9	7.0	92.7	84.4	209.6	692.9	14.7	4.6

Trees	9	10	11	12	13	14	15	16
defoliated-J	23.1	4.6	211.7	36.7	41.1	346.4	10.4	80.2
undefoliated-J	20.6	6.0	2450	38.7	18.9	427.9	7.2	40.0
defoliated-M	25.1	4.2	220.4	41.3	45.9	325.7	13.1	97.0
undefoliated-M	24.2	8.6	261.8	44.9	24.0	446.2	9.1	48.3

where: letters J and M stay for Januar and May results but the numbers for compounds: 1. α-pinene, 2. camphene, 3. β-pinene, 4. myrcene, 5. limonene, 6. β-phellandrene, 7. α-ocimene, 8. α-copaene, 9. bornyl acetate, 10. α-bergamotene, 11. β-caryophyllene, 12. α-humulene, 13. γ-muurolene, 14. germacrene D, 15. β-selinene, 16. β-cadinene

The system is fully described by two factors (principal components) with total eigenvalues of 90% (Fig. 2). Four sets of cases, defoliated and undefoliated trees from January and May, are found to be separated by this two factors. Factor 1 distinguishes defoliated from undefoliated trees but factor 2 stays for differences in climate in January and May.

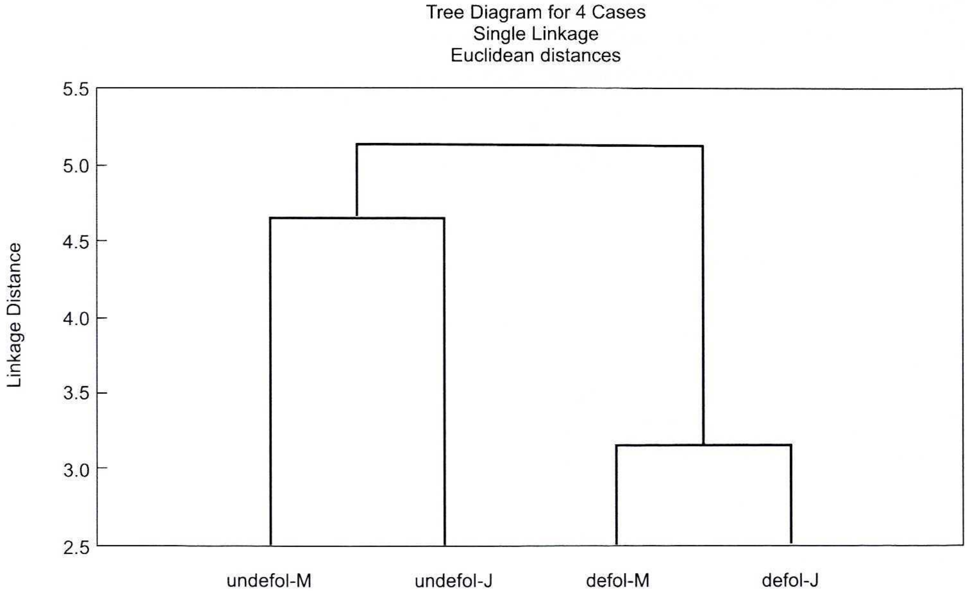

Fig. 1. Cluster analysis of *Pinus caribaea* trees in January and May

Table 2

Eigenvalues and Variances of the Factors

Eigenvalues Extraction: components	Principal			
	Eigenval	% total Variance	Cumul. Eigenval	Cumul. %
1	10.4163	65.102	10.4163	65.102
2	4.77784	29.8615	15.1941	94.9634
3	0.1868	1.16752	15.381	96.131
4	0.04762	0.29762	15.4286	96.4286
5	0.04762	0.29762	15.4762	96.7262
6	0.04762	0.29762	15.5238	97.0238
7	0.04762	0.29762	15.5714	97.3214
8	0.04762	0.29762	15.619	97.619
9	0.04762	0.29762	15.6667	97.9167
10	0.04762	0.29762	15.7143	98.2143
11	0.04762	0.29762	15.7619	98.5119
12	0.04762	0.29762	15.8095	98.8095
13	0.04762	0.29762	15.8571	99.1071
14	0.04762	0.29762	15.9048	99.4048
15	0.04762	0.29762	15.9524	99.7024
16	0.04762	0.29762	16	100

Figure 3 represents a relation between variables and factors, which assign the chemicals taking part in factor 1 responsible for resistance of the trees to ant attacks. Some variables are strongly correlated. The highest loading factors were found for limonene,

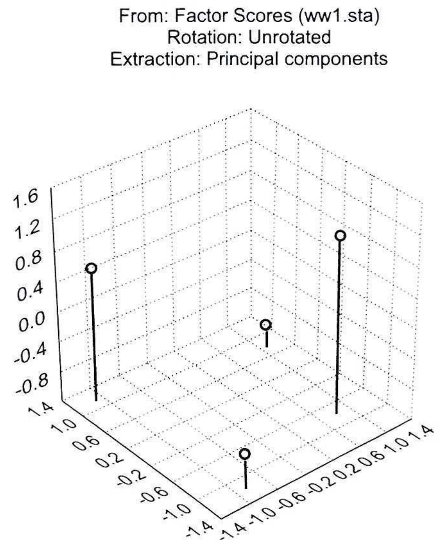

Fig. 2. Factor analysis

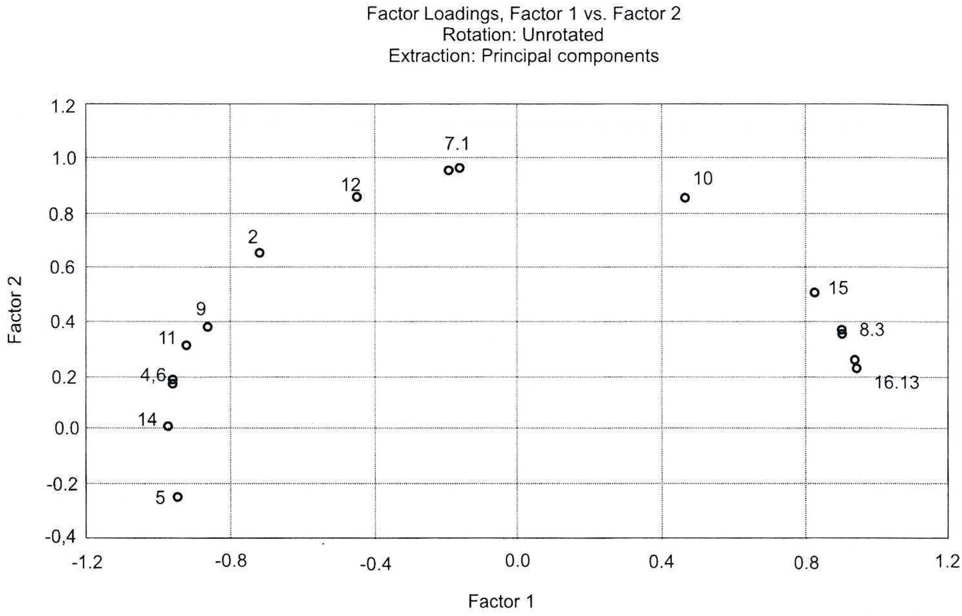

Fig. 3. Factor loadings

germacrene D, myrcene, β-phellandrene, β-caryophyllene (negative correlation) and γ-muurolene, β-cadinene, β-pinene, α-copaene, β-selinene (positive correlation). The compounds mentioned appear in needle waxes at low and high concentrations. Thus a conclusion can be drawn that their specific concentration is responsible for tree-insect interaction.

Barnola et al. (1994) found with usage discrimination analysis that myrcene, β-caryophyllene and α-humulene control in January the behaviour of the insects although the last one was found in both groups of trees defoliated and undefoliated. It seems that the significance of α-humulene in the trees classification is due to its strong correlation with β-caryophyllene.

According to Howard et al. paper (Howard et al. 1988), β-caryophyllene is a growth inhibitor of *Atta cephalotes* fungi. Furthermore, it decreases herbivory of *Lepidoptera* on the plants.

Many compounds are important for insect herbivory on different plant varieties. Physical properties, feed quality and water and tannin concentration is the most important. But the compounds with selective interactions are significant, too. Chemometrics analysis of *Pinus caribaea* and ants *Atta laevigata* system showed that the concentration of terpenes and sesquiterpenes in pine needles are important. Some of them are included in factor 1, which controls the ecological relationships.

IV. REFERENCES

1. Barnola L.F., Hasegawa M., Cedeno A. 1994. Mono- and sesquiterpene variation in *Pinus caribaea* needles and its relationship to *Atta laevigata* herbivory. Biochem. Syst. Ecol., 22: 437–445.
2. Howard J.J., Cazin J., Wiemer D.F. 1988. Toxicity of terpenoid deterrents to the leafcutting ant *Atta cephalotes* and its mutualistic fungus. J. Chem. Ecol., 14: 559–69.
3. Kowalski B.R., Bender C.F. 1972. Pattern recognition. A powerful approach to interpreting chemical data. J. Am. Chem. Soc., 94: 5632–5639.
4. Mardia K.V., Kent J.T., Bibby J.M. 1979. Multivariate Analysis. Academic Press. N.Y.
5. Massart D.L., Kaufman L. 1983. Hierarchical clustering methods. The interpretation of analytical data by use of cluster analysis. Wiley, N.Y.

V. POLISH SUMMARY

CHEMOMETRIA JAKO METODA OCENY ZALEŻNOŚCI EKOLOGICZNYCH W UKŁADZIE ROŚLINA–OWAD

Wiele różnych substancji chemicznych zawartych w roślinach wpływa na zachowanie owadów, pełniąc funkcje repelentów lub atraktantów. Często nie pojedyncze związki, lecz ich mieszaniny o określonym składzie są aktywne semiochemicznie.

Zmienność składu lotnych roślinnych substancji semiochemiczych można oceniać metodami chemometrycznymi w celu znajdywania składników istotnych w ekologicznej interakcji.

Przeprowadzono pełną analizę chemometryczną danych literaturowych dotyczących zmian składu terpenów i seskwiterpenów z igieł sosny *Pinus caribaea* w odniesieniu do żerowania mrówek *Atta laevigata*. Do klasyfikacji danych zastosowano metodę analizy wiązkowej i metodę głównych składników. Dla opisu ponad 90% zmienności składu wystarczyło zastosować dwa faktory opisujące dane dotyczące drzew defoliowanych i niedefoliowanych. Metodą analizy chemometrycznej z ładunków faktorów zidentyfikowano związki chemiczne istotne w oddziaływaniach ekologicznych.