Electronic Journal of Polish Agricultural Universities is the very first Polish scientific journal published exclusively on the Internet, founded on January 1, 1998 by the following agricultural universities and higher schools of agriculture: University of Technology and Agriculture of Bydgoszcz, Agricultural University of Cracow, Agricultural University of Lublin, Agricultural University of Poznan, Higher School of Agriculture and Teacher Training Siedlce, Agricultural University of Szczecin, and Agricultural University of Wroclaw.

ELECTRONIC JOURNAL OF POLISH AGRICULTURAL UNIVERSITIES

2004
Volume 7
Issue 1
Series FORESTRY

TRACZ W. 2004. NEW ANALYTICAL METHODS IN FORESTRY: INTEGRATION OF EXPERT SYSTEMS AND ARTIFICIAL NEURAL NETWORKS WITH GIS **Electronic Journal of Polish Agricultural Universities**, Forestry, Volume 7, Issue 1.
Available Online http://www.ejpau.media.pl

NEW ANALYTICAL METHODS IN FORESTRY: INTEGRATION OF EXPERT SYSTEMS AND ARTIFICIAL NEURAL NETWORKS WITH GIS

Wiktor Tracz
Department of Forest Management, Geomatics and Forest Economics, Warsaw Agricultural University, Poland

ABSTRACT

An effective solution of complex spatial problems requires integration of different models, which describe functionality of nature systems with Geographical Information Systems (GIS). It requires support of typical spatial analyses procedures with results obtained from modelling these problems. Selection of modelling methods, which is used for the solution of a given spatial problem, depends on the complexity of the system being modelled and the degree to which the behaviour of the system is recognized.

Expert Systems (ES) and Artificial Neural Networks (ANN) belong to Artificial Intelligence (AI). AI is the science of creating and developing artificial machines' abilities in such a way that they are similar to human intelligence – they are able to recognize and understand an observed piece of reality, they can define proper goals and are able to search for ways to achieve these goals.

This article focuses on the problem of effectiveness and purpose of ES and ANN integration with GIS in order to rationalize spatial management of multifunctional forestry.

Key words: forest management, spatial analyses, decision support, GIS, Expert Systems, Artificial Neural Networks

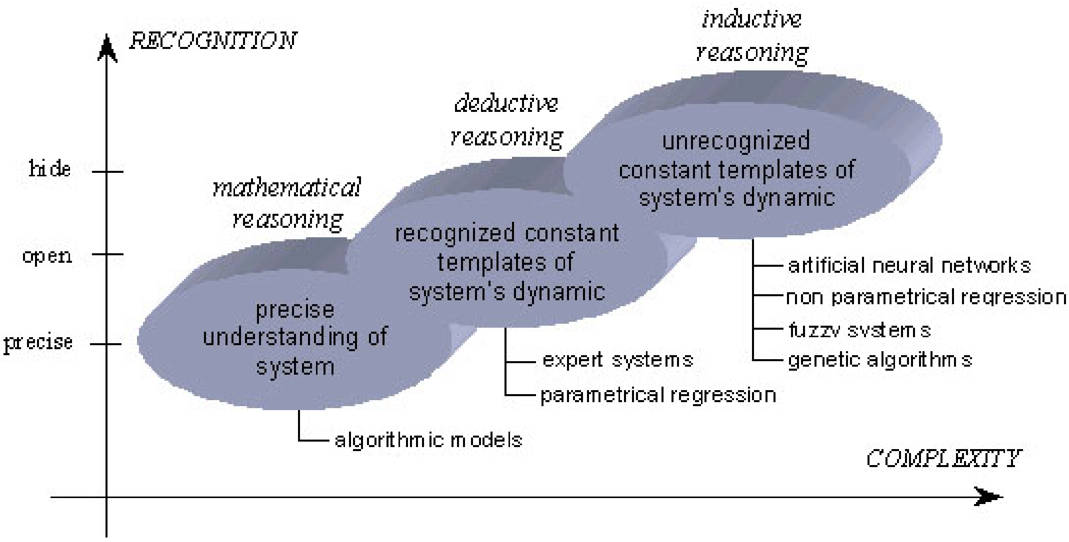
INTRODUCTION

Spatial information is used in forestry in the entire process of matching forest economy to social needs by means of appropriate planning of relationships between productive, protective and recreational functions of the forest. Currently, the most effective tool for supporting spatial objects/events management is Geographical Information Systems (GIS). GIS can be defined as a computer system for gathering, processing and presentation of spatial data. GIS comprises spatial and spatial-accompanying (attribute) data about objects/events chosen in a part of the area which is under the system operation. GIS is based on a general model, which represents reality as separate layers of information.

In order to use GIS, a geographical model of considered object/events has to be created. This geographical model in form of information layers is used to perform spatial analyses in GIS [2, 4, 12]. The complexity of some environmental systems imposes building more advanced models of these systems, which describe systems' objects and processes in a simple – compared to reality – way.

An effective solution of complex (i.e. semi-structured and unstructured) spatial problems requires integration of various models, which describe functionality of nature system with GIS, in other words it requires support of typical spatial analysis processes containing results obtained from models of these systems. The complexity of a system and recognition of its behaviour determine the modelling methods that can be used. The process of model building depends on model functionality knowledge, model complexity, data types, data quality and the amount of data. Zieliński [18] based on Refenes [11] presents the place of modelling methods with regard to their analytical capabilities and inference techniques (Fig. 1).

Fig. 1. Placement of modelling methods



Mathematical models are called 'strong models'. They have strict assumptions and lack free parameters. The nature of the problem being solved must be clearly understood and knowledge about the problem solution, which allows description of modelled system with defined mathematical relationships or equations, must exist. Mathematical models method is undoubtedly one of the most effective ways of problem solving.

Deductive models are used in case of systems for which precise knowledge and clearly defined relationships do not exist. It is possible, however, to find out some constant templates of a system's behaviour based on direct observation of the system. This allows determining general principles of the system's dynamism.

Inductive models are used in cases of more complex systems because that is when the possibility of direct and accurate determination of constant templates of system's behaviour decreases. Relationships between system's variables become implicit. Such problems can be solved using non-parametric methods, which carry out modelling which by induction derives knowledge from data.

Two modelling methods are described in this article: one originates in deductive models and the other in inductive ones. These two methods are Expert Systems (ES) and Artificial Neural Networks (ANN) – techniques being part of Artificial Intelligence (AI). AI is the science of creating and developing artificial machines' abilities in such a way that they are similar to human intelligence – the ability to recognize and understand an observed piece of reality, to define proper goals and to search for ways to achieve these goals.

GIS is a very powerful and widely used technology for spatial analysing and management of spatial systems' elements. Similarly, AI techniques are used in more areas of our life despite a relatively short time of their development and presence among matured tools for data analysing and data processing.

The aim of the research was to prove that, using two chosen problems as an examples, application of ES and ANN in supporting GIS in spatial analyses of forest areas is possible and effective, gives reasonable results and can be successfully used in forestry.

METHOD

Methodology of the research consists of four stages.

1. In the first stage, which includes a bibliography review, a wide analysis of GIS, ES and ANN was done in order to justify the thesis that knowledge of the area of the above-mentioned tools of Information Technology allows their integration and building of new (in the sense of quality) dedicated systems, which effectively support processing of data describing forest phenomena.
2. Spatial problems characterised by their features hindering the building of mathematical models were identified <*footnote 1*>.
3. Review and analysis of software for ES and ANN were performed.
4. Verification of new informatics systems, which were built as the result of integration AI methods with GIS by means of testing empirical examples.

Many researches have been dedicated to GIS applications, and numerous handbooks were published. Rich publications include: [1, 2, 3, 4, 6, 16, 17]. These publication discuss: theoretical bases of GIS, problems of gathering, verification and entering of data, data accuracy, data processing methods, spatial analyses in GIS, results presentation, GIS software, etc. There are many publications regarding GIS together with other techniques in natural environmental study including forestry as well.

The ability to exchange data in different formats is the most essential feature of GIS software, taking into consideration the integration of GIS with ES and ANN. Some of the most popular data formats are recognizable by almost all software irrespective of origin as the result of intensive development of informatics methods.

Expert System can be defined as a computer system which performs complicated tasks in a narrow domain with high intellectual requirements and does this as good as a human expert in this domain.

ES have such properties as integration, prediction, diagnosis and ANN, which carry out classification, recognition, optimisation and association and can be a potential tool for modelling complicated systems and supporting GIS in managing complex forest problems.

ES are used when information (knowledge) about a problem being under consideration is uncertain or incomplete and in cases of semi-structured and unstructured problems, for which a mathematical model of the problem cannot be created [5, 9]. Most important advantages of ES are a natural way of communication with the user and easiness of understanding the way in which ES found the problem's solution.

Artificial Neural Network is a system of simple elements called neurons, connected to each other, which process information. The principle of ANN working is based on a simulation of human brain cell processes by a computer program.

The most important advantage of ANN is its ability to learn. Thanks to this ability and a peculiar structure, ANN demonstrates some capabilities similar to human mental processes, in particular [10]:

- ability to process incomplete data;
- possibility to generate approximate results;
- high tolerance to errors and damages;
- possibility to process fuzzy, chaotic (incoherent) and even inconsistent information.

ANN can learn from empirical data describing a given problem.

The detailed analysis of GIS, ES and ANN features that was undertaken by Tracz [13] proves the thesis about the possibility of integration of the above mentioned informatics tools to the form of new informatics systems in the sense of better functionality.

The main fields of interest in the research presented were Expert Systems and Artificial Neural Networks, which support decision-making processes in management of some forest problems. These particular groups of problems have the following features hindering mathematical modelling:

- semi- or unstructuredness;
- non-linear change in time;
- incomplete knowledge about the system's functionality;
- uncertainty of some inventory data;
- often incompleteness of data caused by high inventory costs;
- often only symbolic (non-numerical) type of data;
- lack of scientific research on the system's functionality.

Several problems, whose features indicate the purposefulness of using ES and ANN for supporting GIS in spatial analyses, were selected based on a preliminary detailed analysis of using GIS in State Forest [8]. The results of the analysis (presented in chapter 'Results') proves the thesis about the presence of spatial problems that can be effectively solved by means of integration of the above mentioned methods.

The ES software can be grouped in 5 groups: algorithmic languages, artificial intelligence languages, expert system languages, specific programs that make system realisation easier, and shall systems. Precisely defined goals, which an expert system will realise supporting GIS, may be the general guide to choosing proper software for the realisation of defined expert system.

The ANN software was chosen based on analysis of 5, considered most advanced ANN simulators. The criteria of the software evaluation were: its abilities, modernity, degree of advance, detailed description and help on the software, accessibility on the Polish market. Each software was evaluated from two points of view. First, as a software and second, as a tool for ANN realisation. The features of good ANN simulator can be described as follows:

- many architectures and training methods;
- reliable and adequate evaluation criterions of trained neural model;
- ability of training data modelling;
- ability of creating ANN hybrid structures.

Tracz [13] presents the review and detailed description of ES and ANN software features. Choosing a software does not affect potential advantages of integrating ES and ANN with GIS. More software that realise ANN and ES methods have similar functionality. Rule-based expert system and artificial neural network built in Multilayer Perceptron (MLP) architecture were considered in this research.

The stage of methodology, that concerns verification of new informatics systems, which were built as the result of integration AI methods with GIS by means of testing empirical examples, is presented in chapter 'Results'.

RESULTS

Several groups of problems were selected from a set of problems noted by GIS user in the National Forest Holding - State Forests (Polish: Państwowego Gospodarstwa Leśnego Lasy Państwowe) – The State Forests in short. These problems possess features, which show purposefulness of using GIS, supported by models created with ES and ANN for their solving. These groups of problems are:

- changes of the spatial structure of the areas adjacent to the forest;
- forecasting of insect population growth;
- forecasting of seed harvest.

When considering the problem of spatial structure changes the problem of field-forest border configuration was the main focus. It had been studied and it was recognized to be currently one of the most important subjects of natural areas configuration in the country.

The problem of forecasting a harmful insects population, the solution of which requires using time series of observations, has been chosen because of its significance in the forestry economy [7].

The two above-mentioned problems were used to test integration AI methods with GIS.

Integration ES with GIS

The verification of ES and GIS integration possibilities was based on a spatial problem of field-forest border configuration [14]. The problem was an optimisation of field-forest border shape while doing afforestation and elimination of enclaves and semi-enclaves inconvenient for management <*footnote 2*>.

Nature areas configuration incorporates many multi-criteria decisions. Multi-criteria problems are difficult to formalize, particularly for such a spatially complex system as landscape.

In order to test integration of ES and GIS, an initial simplistic presumption was made, i.e. the main purpose of the area under consideration (land use category) was wildlife management and the most important animal species for which landscape structure would be optimised was forest roe deer (*Capreolus capreolus*). Field-forest border shape is considered optimal when landscape structure created as result of field-forest border configuration is the most favourable for forest roe deer population and condition.

A prototype of the system for supporting decision-making process of field-forest border configuration optimisation (called the System for short) has module architecture. In general the prototype of the System consists of two main modules: a geographical information system (the GIS) and an expert system (the ES). The System is supported with State Forests' IT System (Polish: System Informatyczny Lasów Państwowych – SILP), Forest Numerical Map (Polish: Leśna Mapa Numeryczna – LMN) and other databases, which deliver e.g. data describing areas outside the State Forests borders, land registration data, etc.

The GIS module is an aggregation of LMN (or a part of it) and GIS for rural district in case when the district area is an area where field-forest border is designed. User interface of the GIS is user interface of System and is fitted in to the needs of the System's users.

Expert system, which is invisible for users, closely co-operates with the GIS module. The ES evaluates shape and spatial structure of landscape by means of using rules, facts (which form knowledge base of the ES) and parameters of landscape.

The whole working cycle of the System consists of three stages. First, analysis of current land use and choosing site for afforestation. Second, change of the category of chosen site to "forest" category. Third, analysis and evaluation of new landscape structure with regard to the chosen land use category.

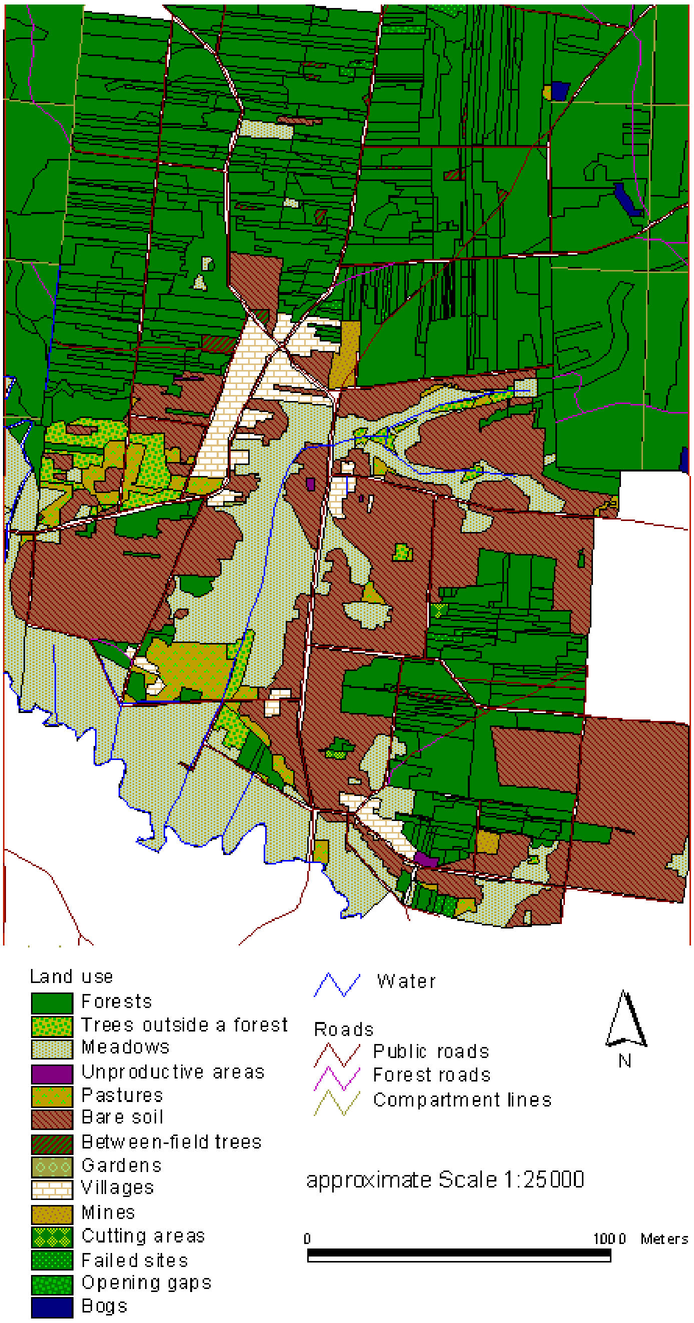
A lands of Wólka Przedmieście village in Wasilków rural commune (Podlaskie Province, Poland) was the object for evaluation of the System prototype. Village lands in this area border on the Dojlidy forest district and private forests. The area was chosen because the Dojlidy forest district possesses Forest Numerical Maps made in accordance with the State Forests Standard. Significant reasons for choosing this area were also: socio-economic characteristic of the region and low soil quality that gives more chances for afforestation of farmlands.

Field-forest border design was optimised to make the population of forest roe deer bigger and to make its condition in the given area better. Knowledge about the considered problem was acquired from domain literature and from consultations with a wildlife management expert.

The System prototype covers all existing guidelines regarding field-forest border design and significant suggestions and recommendations of a classifier – a person who carries out the project of field-forest border design. The System supplied with appropriate geometrical and attributive data can be used to support field-forest border configuration in any part of the country except mountain areas where other principles are crucial.

In order to assess its completeness and effectiveness, the System was demonstrated to the classifier. The GIS module was recognized as adequate, very helpful and effective in the classifier's job. In order to assess effectiveness of the expert system module, evaluation results (suggested by the ES and given by a human expert) of three landscape structures were compared. The results regarding the best landscape structure for forest roe deer life conditions were the same for both the ES and the human expert. The landscape structure presented on Fig. 2 was recognized as the best for forest roe deer population condition.

Fig. 2. The landscape structure chosen by ES as the optimal for forest roe deer life condition



Integration ANN with GIS

Evaluation of ANN and GIS integration functionality was performed in the frame of an empirical model regarding spatial visualisation of harmful insects occurrence and forecasting of insect number change in time and space [15].

The empirical example was based on real data of num moth (*Lymantria monacha L.*) butterflies catch to pheromone traps. Time series of num moth observation covers 5 years. The prototype of the system of num moth population forecasting has module architecture. Two main modules of suggested architecture are geographical information system (the GIS) and artificial neural networks (the ANN). They are supported by the Informatics System of State Forests, the Forest Numerical Map, meteorological data and insect monitoring data.

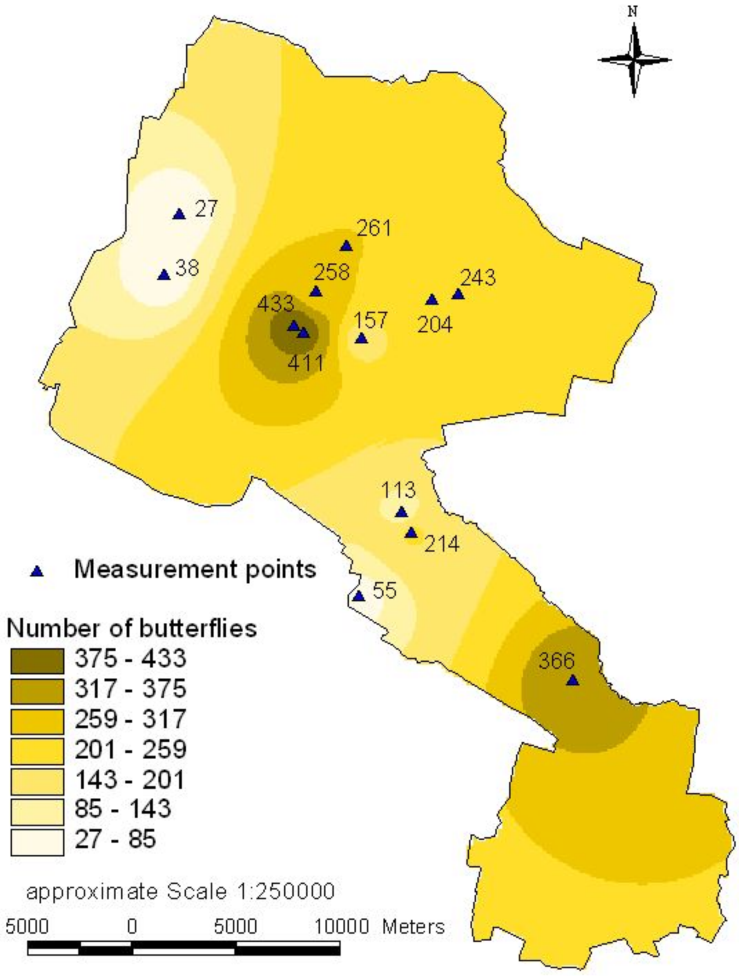
The ANN was used to forecast insect population in places being under consideration. The ANN was built as Multilayer Perceptron (MLP) and back propagation algorithm was used to train the net. As the result the insect population forecasting model was created by the ANN. The indexes available in the software used for ANN simulation were used to assess quality and fitness of the model. Tracz [13, 15] presents detailed analysis and assessment of trained ANN model. The analysis showed that the forecasting model was reliable.

The test object where empirical data was gathered was Skarżysko-Kamienna forest district. The forest district possesses GIS, which was built in accordance with the State Forests Standard.

The forecasted amount of the num moth concerns only the stands where pheromone traps were placed. In order to know the insect number in other stands the values forecasted by the ANN were interpolated in the GIS. As the result of the interpolation the information layer of num moth butterflies occurrence probability in the analysed area was achieved. Spatial distribution of ANN-forecas numbers and interpolated by GIS is presented on Fig. 3.

Evaluation of forecasting relevance shows that the difference between real and forecasted number of insect was equal 7.27%. Such error was recognized by forest protection scientists as minor and showing satisfactory accuracy of ANN forecasting.

Fig. 3. Spatial distribution of num moth butterflies – the result of interpolation forecasted by ANN numbers



CONCLUSIONS

Analysis of GIS user's needs in forestry management shows existence of complex problems. These problems can be solved using advanced analytical methods (e.g. such AI methods as ES and ANN) that support GIS in spatial analyses.

Usefulness and effectiveness of ES were proved based on the field-forest border configuration system example – a problem considered to be important in landscape planning. A properly built and functioning prototype of information system, which is dedicated to landscape structure analyses can be used in scientific research and in landscape design practice. ES is recommended for problems which are relatively well understood but because of their complexity other modelling methods cannot be applied.

ANN is an appropriate method in case a large amount of data (very often time series of data) has to be processed in order to model environmental phenomena. Usefulness and effectiveness of ANN were proved based on insect population forecasting system example.

Results of the research on integrating ES and ANN with GIS prove that functional integration in the sense of good understanding of a spatial problem and properly addressing functions and data for GIS and modelling module, is possible. Technical integration, which is understood as a higher effectiveness of using various software in one coherent system – application dedicated to chosen problem solving, is also possible.

Solutions (empirical examples) presented here are methodical templates, which can be used in other environmental conditions. The designed prototypes of systems could be fully functional applications used on various organisational levels of the State Forests management. The prototype of the system which supports field-forest border configuration could also be used by local governments.

REFERENCES

5. Aronoff S. 1989. Geographic Information Systems. A Management Perspective. WDL Publication. Ottawa.
6. Burrough P. 1986. Principles of Geographical Information Systems for Land Resources Assessment. Oxford University Press. New York.
7. Ciołkosz A. i inni. 1990. System informacyjny o środowisku przyrodniczym. [Natural environment information systems]. Wydawnictwo SGGW-AR. Warszawa [in Polish].
8. Gaździcki J. 1990. Systemy informacji przestrzennej. [Spatial information systems]. PPWK. Warszawa-Wrocław [in Polish].
9. Ignizio J. 1991. Introduction to expert systems. McGraw-Hill Inc.
10. Maguir D., Goodchild M., Rhind D. 1991. Geographical Information Systems: Principles and Applications. Wiley, New York.
11. Mozgawa J. i inni. 2001. Analiza przestrzenna związków przyczynowo-skutkowych pomiędzy stanem zdrowotnym drzewostanów a czynnikami szkodotwórczymi – na poziomie nadleśnictwa. [Spatial analyses of forest damage factors]. Dokumentacja na potrzeby DGLP. Maszynopis [in Polish].
12. Mozgawa J. i inni. 2002. Analizy przestrzenne, optymalizacja i symulacje w zarządzaniu Lasami Państwowymi z uwzględnieniem standardów leśnych map numerycznych na poziomie nadleśnictwa, RDLP i DGLP. [Spatial analyses, optimisation and spatial simulation in Polish State Forests management make allowance for the Standard of the Forest Numerical Map]. Dokumentacja na potrzeby DGLP. Maszynopis [in Polish].
13. Mulawka J. 1996. Systemy ekspertowe. [Expert systems]. WNT. Warszawa [in Polish].
14. Openshaw O., Openshaw Ch., 1997. Artificial Intelligence in Geography. John Wiley and Sons.
15. Refenes A. (red.), 1995. Neural Networks in the Capital Markets. John Wiley and Sons. Chichester.
16. Tomlin C. 1989. Geographical Information Systems. Prentice Hall.
17. Tracz W. 2002. Analiza obszarów leśnych w Systemach Informacji Przestrzennej wspomaganych Systemami Ekspertowymi i Sztucznymi Sieciami Neuronowymi. [Using Artificial Neural Networks and Expert Systems for supporting GIS in spatial analysis of forestry areas management]. PhD thesis. Praca doktorska. Wydział Leśny SGGW [in Polish].
18. Tracz W. 2003a. GIS supported with Expert Systems in optimisation of field-forest border shape. Folia Forestalia Polonica nr 45, 2003:47-55.
19. Tracz W. 2003b. Wykorzystanie Sieci Neuronowych wspomagających SIP w analizach przestrzennych w leśnictwie. [Use of GIS supported by Artificial Neural Networks in spatial analyses in forestry]. Sylwan nr 12, 2003:36-44 [in Polish].
20. Werner P. 1992. Wprowadzenie do geograficznych systemów informacyjnych. [Introduction to Geographical Information Systems]. Wydział Geografii i Studiów Regionalnych. Uniwersytet Warszawski Warszawa [in Polish].
21. Widacki W. (red.). 1993. GIS for Environment. Conference on Geographical Information Systems in Environmental Studies. Instytut Geografii Uniwersytetu Jagiellońskiego. Kraków.
22. Zieliński J. 2000. Inteligentne systemy w zarządzaniu. Teoria i praktyka. [Intelligent systems in management. Theory and practice.]. WNT. Warszawa [in Polish].

FOOTNOTES

1. This stage of the research was based on the results of two research projects, which were done in the Department of Forest Management, Geomatics and Forest Economics of Warsaw Agricultural University on order of General Directorate of State Forests (Polish: Dyrekcja Generalna Lasów Państwowych – DGLP).
2. The research concerning this problem was financed by Polish State Committee for Scientific Research (Polish: KBN) along 6 P06L 03521 project.

Wiktor Tracz
Department of Forest Management, Geomatics and Forest Economics
Warsaw Agricultural University
ul. Nowoursynowska 159, 02-776 Warsaw, Poland
tel. +48-22-8439061 ext. 38018
e-mail: wiki@witch.sggw.waw.pl