Modelling Non-Industrial Private Forest Landowners' Strategic Decision Making by Using Logistic Regression and Neural Networks: Case of Predicting the Choice of Forest Taxation Basis

Mauno Pesonen, Petri Räsänen and Arto Kettunen

Pesonen, M., Räsänen, P. & Kettunen, A. 1995. Modelling non-industrial private forest landowners' strategic decision making by using logistic regression and neural networks: Case of predicting the choice of forest taxation basis. Silva Fennica 29(2): 171–186.

In this study, logistic regression and neural networks were used to predict non-industrial private forest (NIPF) landowners' choice of forest taxation basis. The main frame of reference of the study was the Finnish capital taxation reform of 1993. As a consequence of the reform, landowners were required to choose whether to be taxed according to site-productivity or realized-income during the coming transition period of thirteen years.

The most important factor affecting the landowners' choice of taxation basis was the harvest rate during the transition period, i.e. the chosen timber management strategy. Furthermore, the estimated personal marginal tax rate and the intention to cut timber during next three years affected the choice. The descriptive landowner variables did not have any marked effect on the choice of forest taxation basis.

On average, logistic regression predicted 71 % of the choices correctly; the corresponding figure for neural networks was 63 %. In both methods, the choice of site-productivity taxation was predicted more accurately than the choice of realized-income taxation. An increase in the number of model variables did not significantly improve the results of neural networks and logistic regression.

Keywords forest taxation, logistic regression, neural networks, non-industrial private forest landowner, timber management strategies.
Authors' address Finnish Forest Research Institute, P.O. Box 18, FIN-01301 Vantaa, Finland **Fax** +358 0 8570 5809 **E-mail** mauno.pesonen@metla.fi
Accepted September 25, 1995

1 Introduction

1.1 Strategic Decision Making

Most non-industrial private forest (NIPF) landowners have long-term perspectives in regard to their strategic view of forest management (Lönnsted 1989). It is important to understand the strategic decisions of NIPF landowners for several reasons: e.g. predictions of the timber supply from private forests for investment plans by forest industries (Lönnstedt and Roos 1993).

Strategic planning operates on future production possibilities; the starting point in planning is in the variability of the factors of production and their allocation (e.g. Kast and Rosenzweig 1974). When applied to NIPF management planning, the strategic view includes the production of alternative, strategic-level programmes for timber production and silviculture. Timber management covers a range of strategies from "no cuttings at all" to "maximum cuttings within the limits of timber production possibilities". Timber management strategies can be described in terms of intensity and recurrence of cuttings, for instance.

Many studies (e.g Wardle 1965, Kilkki 1968, Ware and Clutter 1971, Kangas and Pukkala 1992, Siitonen 1983, Johnson et al. 1986, Jonsson et al. 1993, Pukkala and Kangas 1993) have been done on the subject of strategic forest management planning. Strategic-level decision making and decision processes have been studied by Lönnstedt and Törnqvist (1990), Kajanus (1992), Pukkala and Kangas (1993). Forest taxation, as part of forest management planning, has received little attention.

1.2 Forest Taxation Reform

In 1993, the Finnish forest taxation system underwent a reform, when site-productivity taxation (SPT) was replaced by realized-income taxation (RIT). The Finnish forest taxation reform includes a 13-year transition period for non-industrial private forest (NIPF) landowners.

The choice of forest taxation basis is part of the strategic decision making of landowners, as is also the choice of timber management strategy (Pesonen 1995). In the spring of 1994, landowners were required to choose whether to be taxed according to SPT or RIT for the next 13 years. During this transition period of 13 years (1993–2005), landowners choosing SPT will be able to realize their accumulated timber growth which has already been taxed once prior to the taxation reform.

SPT is based on the estimated taxable income, i.e. the estimated value of the mean annual increment according to the site's soil productivity (Laki maatilatalouden ... 1990). Under this system, the estimated taxable income from forestry is added to the NIPF landowner's non-forestry income, and the final and actual tax to be paid annually depends on the landowner's personal marginal tax rate. The new RIT system is based on the individual landowner's annual timber sales revenues. The net timber income is taxed applying a uniform and constant tax rate, which in 1993 was 25 % (Tuloverolaki 1992).

The most important factor affecting the individual NIPF landowner's optimal choice of taxation basis is the harvest rate during the transition period, i.e. the chosen timber management strategy. The estimated value of annual increment under SPT, the landowner's personal marginal tax rate, and the uniform capital tax rate were also significant factors influencing the optimal forest taxation basis (Ovaskainen et al. 1992, Pesonen and Räsänen 1993). In order to assist NIPF landowners in making their choice, Pesonen and Räsänen (1993) formulated a model for the optimal forest taxation basis. In the model, the optimal choice depends on the landowner's marginal tax rate and the relation between the chosen timber management strategy and the volume of taxable increment (m^3/ha/a), i.e. the estimated growth under SPT.

1.3 Logistic Regression and Neural Networks

Logistic regression has been rarely used in modelling strategic decision-making problems. Instead, logistic regression models have been widely used in solving other kinds of problems. Royer (1987), for example, has modelled the reforestation behaviour of NIPF landowners, Heliövaara et al. (1991) have predicted the distribution of bark beetles using climatic variables, and

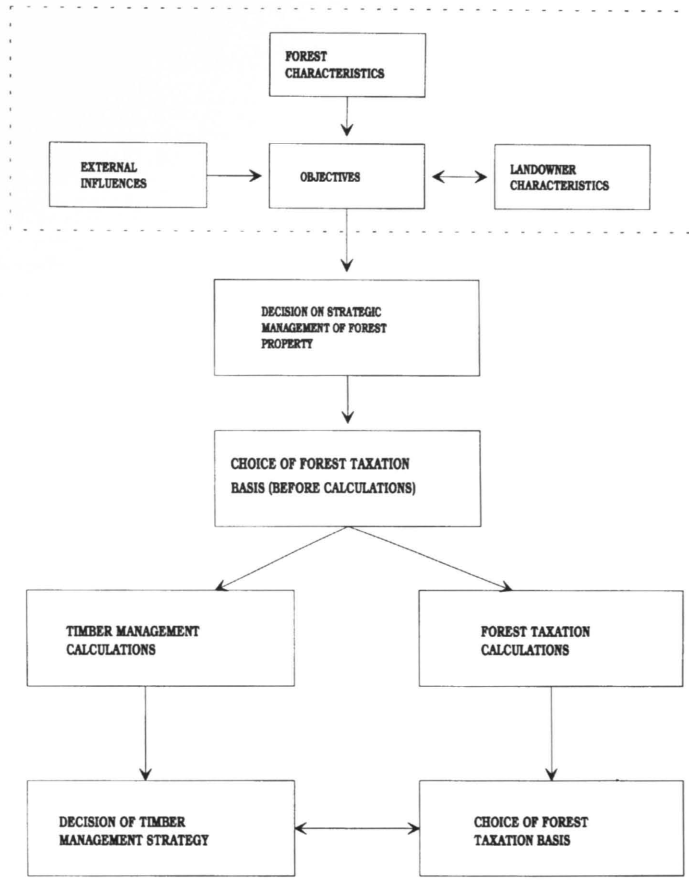

Fig. 1. Frame of reference applied in this study.

Uusitalo (1993) has predicted the timber quality of Scots pine.

Methods in the field of machine learning offer new approaches to retrieving knowledge from ill-defined, noisy domains. Many machine learning techniques are non-parametric in character and capable of dealing with quantitative as well as qualitative variables with linear and nonlinear dependencies. Often they are noise-tolerant and nonsensitive to fixed hypotheses (Guan and Gertner 1991).

The basic idea in neural networks is to simulate and understand the processes of a nervous system. In forest science, neural networks have been used, for example, in estimating non-optimality losses resulting from harvesting decisions (Lämås et al. 1991) and in estimating forest stand characteristics based on satellite-based remote sensing (Feycting et al. 1991). Nikula and Väkevä (1991) have modelled the risk of moose-browsing in forest plantations and Guan and Gertner (1991) have modelled red pine survival. Furthermore, neural networks have been applied to forecasting recreation in wilderness areas (Pattie

1992) and to predicting processing parameters in particleboard manufacturing (Cook et al. 1991). As yet, neural networks have not been applied in modelling strategic decision-making problems.

1.4 Aim of the Study

The aim of this study is to 1) predict NIPF landowners' choices using logistic regression and neural networks, 2) compare the performance of the two methods in predicting NIPF landowners' choices of forest taxation basis, and 3) clarify the factors affecting landowners' actual choices between site-productivity taxation and realized-income taxation.

In modelling the choice of forest taxation basis, logistic regression is used because of the binary outcome of the choices. The frame of references of the study is presented in Fig. 1.

2 Methods

2.1 Logistic Regression

The form of the logistic regression model (1) was

$$\mathrm{Log}\left(\frac{\pi}{1-\pi}\right) = \mathrm{Log}(O) = \alpha + \beta_1 x_1 + \beta_2 x_2 + \beta_n x_n \quad (1)$$

where

π	conditional probability of choosing site-productivity taxation
$1-\pi$	conditional probability of choosing realized-income taxation
$\alpha, \beta_1,\ldots,\beta_n$	parameters of logistic regression
$x_1,x_2,\ldots,x_n$	independent variables
O	conditional odds of choosing site-productivity taxation
n	number of independent variables.

The betas ($\beta_1,\ldots\beta_n$) represented the change in the log-odds due to unit increments in the value of the dependent variables. Moreover, the exponential coefficient $\exp(\beta_i)$ was the coefficient of the odds ($\pi / (1-\pi)$) for a one-unit increase in the *i*th predictor, and 100 [$\exp(\beta_i) - 1$] was the estimated percentage change in the odds for a one-unit increase in the *i*th predictor. The method of maximum likelihood was used to compute estimates of the parameters $\alpha,\beta_1,\ldots\beta_n$ (Demaris 1992, BMDP ... 1992).

2.2 Neural Networks

2.2.1 Principle of Neural Networks

A neural network, or a parallel distributed processing model, is a system consisting of a number of simple, highly interconnected processing elements. Neural network models are non-parametric in character and make inferior assumptions than classical statistical methods about independent variables (Cook et al. 1991). There are several types of neural networks: e.g. back-propagation networks, self-organizing maps, and hopfield networks. The networks differ from one another in, for example, network size, input or output type, and the training method applied (Bailey and Thompson 1990).

The basic unit in neural networks is the processing element (PE), analogous to the biological neuron (Fig. 2) (Rumelhart and McClelland 1986). A PE receives inputs from its neighbours and, as a function of the inputs it receives and through the transition function, the PE produces an output value, which it then sends to its neighbours (Fig. 2). The transition function can be either a threshold or a stochastic function. The network stores knowledge implicitly in a set of connection weights; e.g. if the weight between units u_i and u_j is a positive number, unit u_i excites unit u_j and if the weight is negative, unit u_i inhibits unit u_j. The absolute value of w_{ij} specifies the strength of the connection (Rumelhart and McClelland 1986).

Neural networks usually include three kinds of processing elements (or nodes): input, output and hidden nodes. Input and output nodes connect the network to the outside world, e.g. to the process parameters being measured (Cook et al. 1991). A series of input nodes comprises an input layer while output nodes make up the output layer. Between these layers of nodes there may be one or more hidden layers, which, in turn, are connected to the other layers. The way the neural network changes the weights between

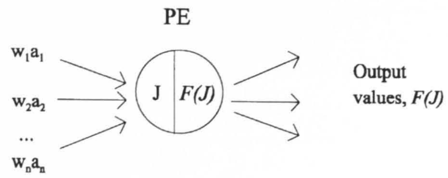

Fig. 2. The basic unit in a neural network ($w_1..w_n$ = weighting factors, $a_1..a_n$ = input values of the processing element (PE), $J = \sum w_i a_i$ and $F(J)$ = transformation function).

the processing elements is referred to as the learning method. An example of learning methods applied to neural networks is the generalized delta rule method used in back-propagation networks (Rumelhart and McClelland 1986).

2.2.2 Back-Propagation Algorithm

A back-propagation-type of neural network was used in this study. This kind of a network first uses input data sets to produce a random output, and then compares this with the observed output. The differences between the output predicted by the network and the observed output are called error signals. Thus, the back-propagation method belongs to the group of supervised learning methods, because the correct result of each input data set is required to control the learning process (Bailey and Thomson 1990).

The main idea in a back-propagation network is to minimize the sum of errors by manufacturing the weights of the connections. The weight changes are first made for all connections feeding into the final layer, and once this has been done, the error signals for all units in the previous layer are computed. This propagates the errors back one layer, and the same process is repeated for every layer. The learning process continues until the network finds a single set of weights satisfying the input/output pairs presented to the network (Rumelhart and McClelland 1986, Cook and Wolfe 1991).

The neural network system is inherently parallel in that many units can execute their computations concurrently (Rumelhart and McClelland 1986). Usually, the primary data are divided into two parts: a training set, which is used in teaching the neural network, and a test set, which is used in evaluating the results of the network. In testing, each observation of the test data is fed in the network and the error between the predicted and the original value is then calculated.

The back-propagation network used in this study was composed of three layers. The input layer consisted of independent variables scaled between 0 and 1. The hidden layer consisted of 11 nodes, which was the best number of nodes tested in this study, and the network had a single output node: the estimated probability of choosing site-productivity taxation. The way the network was constructed, the default value of selections of cases learnt was 500 000 unless the network converged before that. In the calculations, software called NeuroShell (NeuroShell 1989) was used.

3 Material

3.1 Data and Choices of Timber Management Strategy and Forest Taxation Basis

The data were collected from the area under the jurisdiction of the Pohjois-Savo forestry board, in eastern Finland. The basic information about the forest holdings consisted of the forestry plans made according to the TASO planning system managed by the Forestry Centre Tapio (Ranta 1991). Descriptive information about landowners, their forestry property and forestry goals were collected by means of a two-phase mail inquiry (Pesonen 1995). The total sample consisted of 757 forest holdings.

NIPF landowners were asked as to their intended choice of forest taxation basis in two phases (Fig. 3). In addition to the intended choice of forest taxation basis, they were presented with the arguments for a particular choice, detailed information about NIPF landowners' holdings (e.g. the number of taxation cubic meters under SPT) and asked about landowner's taxable income in the first inquiry. The number of acceptable answers received was 413.

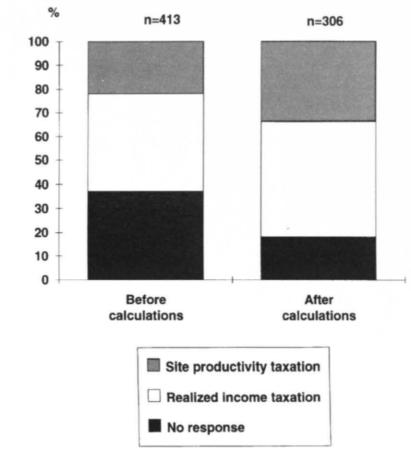

Fig. 3. The choices of forest taxation basis made by the forest owners in the two mail questionaries.

Following the first inquiry, five timber management strategies were calculated using the MELA system (Kilkki and Siitonen 1976, Siitonen 1983, Siitonen 1993) for each NIPF landowner for a planning period of 20 years. The applied strategies were (Pesonen 1995):

1 NO CUTTINGS (Total abstaining from cuttings)
2 SAVING (Utilizing approx. half of the sustained allowable cut)
3 SUSTAINABILITY (Practising forestry on sustained yield basis)
4 FINANCE (Utilizing majority of the allowable cut during the first 10-year period)
5 MAX CUTTINGS (Instantly utilizing the total allowable cut).

The forest-holding-level development of, for example, removals, growth and total volume of growing stock were presented to the landowners

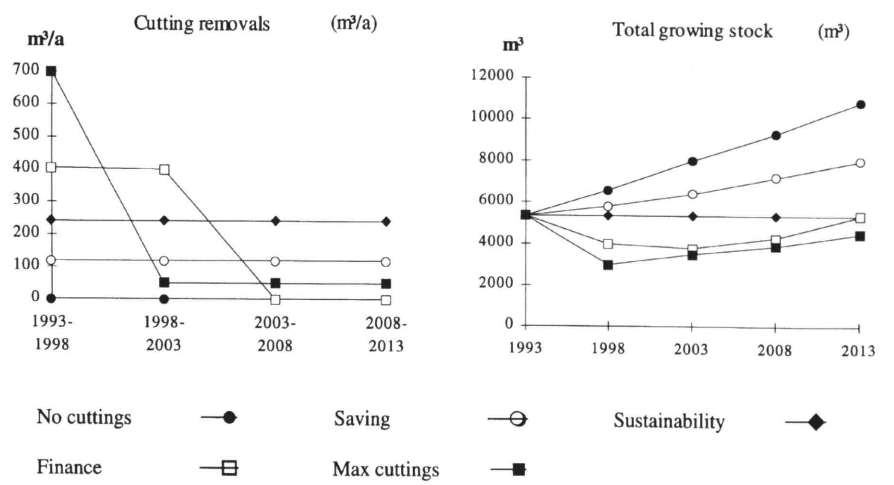

Fig. 4. Alternative timber management strategies (Pesonen 1995) described as the development of the removals and the total growing stock during the planning period (an example of calculations for each forest owner, representing a sample case of the forest holdings).

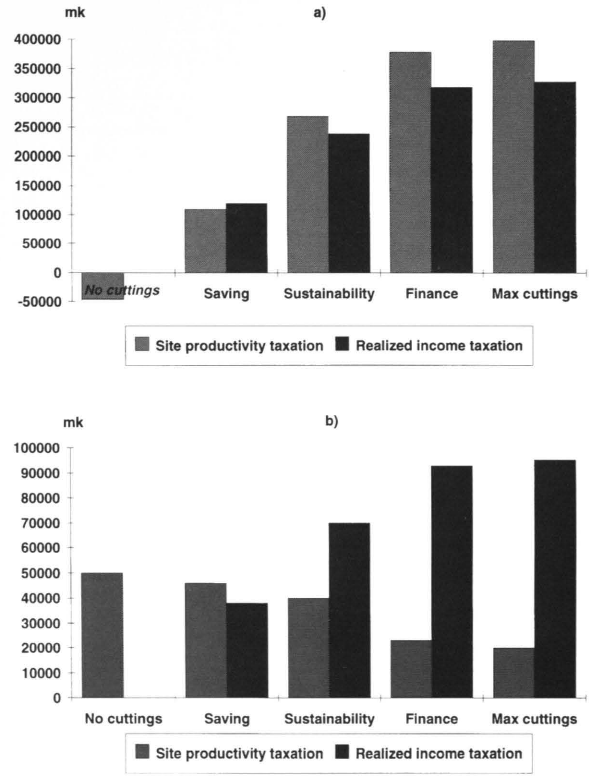

Fig. 5. Profit (a) and payble taxes (b) for the transition period (an example of calculations for each forest owner, representing a sample case of the forest holdings).

(Fig. 4). The landowners were then asked to prioritize alternative timber management strategies by using the Analytic Hierarchy Process (AHP) (Saaty 1977, 1980). The highest priority obtained from AHP represented the preferred timber management strategy (Pesonen 1995).

The profits and the amount of payable taxes for the transition period in each timber management strategy were illustrated for the NIPF landowners (Fig. 5). According to the example presented below, if the landowner intended to cut less than the SAVING strategy, RIT would have been the optimal choice, but if he intended to cut according to sustainability or more, the optimal choice would have been SPT (Fig. 5). Based on these illustrations of comparison, the landowners were asked to re-assess the question of forest taxation basis (Fig. 3). After the two inquiries, the final sample consisted of 306 NIPF landowners.

3.2 Variables Used in the Study

In the analyses, the dependent variable was the choice of forest taxation basis both before and after the profitability calculations. In the first inquiry (before landowners were shown the comparison pictures), the landowners had to make the choice without the benefit of information about the economic consequences of their choice of forest taxation basis in alternative timber management strategies. It can be assumed that these choices were more or less based on personal conceptions of the alternative forest taxation bases. In the second phase (after the profitability calculations were shown to the landowners), the NIPF owners were able to use the comparison pictures of the two forest taxation bases when making the choice between the two taxation systems. In the second phase, it can be assumed that these choices were based more on economic and rational facts than the choices made before being shown the calculations.

Timber management strategy has been found to be a significant variable in modelling the optimal choice of forest taxation basis (Pesonen and Räsänen 1993). Therefore, only those NIPF landowners with information on the preferred timber management strategy (n = 306), were included in the analyses. The analyses were completed only for non-industrial private forest landowners; corporate ownership was not included, because the taxation system and the decision making process in corporately-owned holdings are somewhat different to those applied in privately owned forests. Furthermore, the "no-response" observations were ignored because of the heterogeneity of the group. The number of "no-response" observations decreased in the second inquiry (Fig. 3). As a result, the number of landowners in modelling the first choice was 147, and in modelling the second choice (after the profitability calculations) it was 193 (Table 1).

Seven independent variables were used to describe both choices of forest taxation basis (Table 1). In earlier studies, it has been found that the future cuttings during the transition period divided by the volume of taxable increment in SPT (here referred as forest taxation index) is the most important factor affecting the optimal forest taxation basis (Ovaskainen et al. 1992, Pesonen and Räsänen 1993). The forest taxation index describes the "harshness" of SPT at varying harvest rates. Furthermore, the landowner's estimated marginal tax rate and his/her intention to cut timber in the near future (an expressed intention to cut or not to cut timber during 1993–1995) influenced the optimal choice (Pesonen and Räsänen 1993). The landowners' personal marginal tax rate under SPT depicted the harshness of their personal taxation level and their intention to cut timber depicted the short-term cutting of timber while the timber management strategy depicted the average cutting of timber during the transition period. In addition, the average commercial cuttings of timber of the past 5-year period were assumed to affect the future cuttings of timber and, thereby, also the choices of the forest taxation basis.

The age of the landowner has been found to be a significant variable describing the forestry behaviour of NIPF landowners, e.g. the life cycle harvest of NIPF landowners (Kuuluvainen and Salo 1991). One of the variables discovered to affect the timber management strategy choice was the area of forest land owned by the person (Pesonen 1995). Ovaskainen et al. (1992) assumed also that the forest area owned could affect the choice. Therefore, according to Ovaskainen et al. (1992), if the forest area is smaller than 15 ha, the landowner is assumed to always choose RIT. Furthermore, in that study, farmers were supposed to choose SPT more frequently than non-farmers, because of the debt burden of farmers and the poor profitability of agriculture (Ovaskainen et al. 1992).

3.3 Cross-Validation and Classification

In the comparison between logistic regression and neural networks, both of the data sets (choices of taxation basis before and after the profitability calculations) were randomly cross-validated into seven training sets and seven test sets. Thus, a total of 14 (= 7 x 2) logistic regression equations and neural networks were obtained from the training sets. Cross-validation was applied in order to increase the reliability of the results. With a single division of the material into training and testing sets, the results obtained depend more on the random effect of the division (e.g. Lämås et al. 1991). In modelling the first choice, the training sets included 100 landowners and, in modelling the second choice, there were 135 landowners.

After cross-validation, the estimated values of the observations in each test set were calculated using the regression equation constructed by using the respective training set. After calculating the probabilities, the test observations were classified according to the actual distribution of the choices obtained from the respective training set (the cut-off point). For instance, if the number of actual choices of RIT was 34 in the training set and the number of choices of SPT was 24, the cut-off point used would be 0.41 (= 24 / (34 + 24)). If the estimated probability of an observation was greater than 0.41, the predicted choice of forest taxation basis was SPT (1), and if the estimated probability was lower than 0.41, the predicted choice was RIT (0). This classification is generally done by using a cut-off point value of 0.5, but if the sample is relatively unbalanced, the cut-off point of 0.5 may never predict a case to category 1 (or zero) (Greene 1993).

Table 1. Summary of the variables used in this study.

	Before calculations	After calculations
Number of observations [1]	147	193
Choice of the site productivity taxation	49 (33 %)	84 (44 %)
Choice of the the realized net income taxation	98 (67 %)	109 (56 %)
The age of the landowners (AGE)	51	51
Minimum-Maximum	22–87	22–87
Standard deviation	13	14
Farmer/Non-farmer, Dummy (FARMER)		
Number of farmers (1)	92 (63 %)	121 (63 %)
Number of non-farmers	55 (37 %)	72 (37 %)
Area of forest land, ha (AREA)	65	67
Minimum-Maximum	8–605	8–605
Standard deviation	69	66
Average commercial cuttings of the past 5-year period, m^3/ha/a (CUTTINGS)	2.4	2.3
Minimum-Maximum	0–13.7	0–13.7
Standard deviation	2.5	2.5
Intention to cut during 1993–95, Dummy (CU9395)		
Intends to cut (1)	125 (85 %)	164 (85 %)
No cuttings	22 (15 %)	29 (15 %)
The estimated marginal tax rate in 1991, % (RATE)	41	39
Minimum-Maximum	0–56	0–56
Standard deviation	13	15
Forest taxation index [2] (INDEX)	1.6	1.7
Minimum-Maximum	0–5.5	0–5.5
Standard deviation	0.9	0.9

1) No-response cases were ignored.
2) Forest taxation index = the timber management strategy of the landowners (m^3/ha/a) divided by the volume of taxable increment in site productivity taxation (m^3/ha/a). The forest taxation index describes how much the landowner intends to cut in relation to the mean annual increment during the transition period.

4 Results

4.1 Factors Affecting the Choice of Forest Taxation Basis

In order to clarify the factors which affected the landowners' choice of forest taxation basis, two

Table 2a. Logistic models for the choices of the forest taxation basis before calculations.

Variable	Coefficient	Coeff./SE
Intercept	–3.385	1.45
Age	–0.025	–1.53
Cu9395	1.930	1.79
Index	1.237	4.26
χ^2 [1)]	7.333	
Df	8	
N	147	

Table 2b. Logistic models for the choices of the forest taxation basis after calculations.

Variable	Coefficient	Coeff./SE
Intercept	–1.769	–2.47
Cu9395	1.337	2.34
Rate	–0.030	–2.55
Index	0.862	3.93
χ^2 [1)]	18.193	
Df	8	
N	193	

[1)] The test is the Hosmer-Lemeshow test because of the use of the real-valued variables. The Hosmer-Lemeshow goodness of fit chi-square test divides the data into 10 cells and compares the observed and predicted frequencies for these cells. The cells are defined using the predicted frequencies (DeMaris1992 , BMDP ... 1992).

logistic regression equations were constructed with seven independent variables (Appendix). Because the results of these analyses were poorer than those of analyses using significant variables, the results are reported using only models with significant variables. The final model consisted of significant variables with P-values lower than 0.20 (BMDP... 1992).

The most significant variable in modelling the choices of the taxation basis was the forest taxation index (Index) and it was the included in both models (Tables 2a and 2b). In modelling the first choice, the other variables of the model were "intention to cut timber during next three years" (Cu9395) and "age of landowner" (Age) (Table 2a). In modelling the second choice,the significant variables included in the model, in addition to forest taxation index, were "marginal taxation rate" (Rate) and "intention to cut timber in the near future" (Cu9395) (Table 2b). The occupational status of the landowner (farmer vs. non-farmer), the area of the forest holding, and the average cuttings of the past 5-year period did not significantly influence the choice (Appendix).

Were the landowner to double the cutting of timber during the transition period, i.e. forest taxation index increases from 1 to 2, the odds of choosing SPT ($\pi / (1 - \pi)$) would increase by more than 200 % in the first-choice model, and over 130 % in the second-choice model. A one-unit increase in the tax rate would decrease the odds by 3.0 % in modelling the second choice. Furthermore, the log-odds were observed, in both models, to be very sensitive to the intention to cut timber during the next three years.

4.2 Comparison between Logistic Regression and Neural Networks

The performance of logistic regression and neural networks was compared by using seven cross-validated data sets. Moreover, comparisons were made between models with seven independent variables (Appendix 1) and with three significant variables (Tables 3 and 4). Because the results of the models with three variables were better both in logistic regression and in neural networks, only these results are reported.

In the first choice (before calculations), logistic regression predicted 72 % of the cases correctly (Table 3), while neural networks predicted only 64 % of landowners' choices correctly. Both methods gave better results in predicting the choices favouring SPT. The proportions of the correctly predicted choices were 75 % in logistic regression, and 69 % in neural networks.

In modelling the second choice (after calculations), the proportion of correctly predicted choices was 7 percentage units higher for logistic regression than for neural networks (Table 4). As was the case with models for the first choice, both methods gave better results in predicting the choice of SPT than the choice of RIT: of the predictions of choices favouring RIT as the second choice, 66 % were correct for neural networks, and as many as 87 % were correct for logistic regression.

Table 3. Summary of the modelling of the choice before calculations. SPT = Site-productivity taxation, RIT = realized-income taxation, N = the average number of the landowners in seven different models, Stdev. = the standard deviation of the number of landowners in different models and % = the average percentage of the correctly modelled choice.

Logistic regression				Neural networks			
Modelled / actual	N	Stdev.	%	Modelled / actual	N	Stdev.	%
SPT / SPT [1)]	12	2.27	**75**	**SPT / SPT** [1)]	11	5.08	**69**
RIT / SPT	4	1.95		RIT / SPT	5	5.46	
SPT / RIT	9	2.75		SPT / RIT	12	6.87	
RIT / RIT	22	2.52	**71**	**RIT / RIT**	19	6.26	**61**
Total	47		**72**		47		**64**

[1)] E.g., the modelled choice was site-productivity taxation and the actual choice was site-productivity taxation.

Table 4. Summary of the modelling of the choice after calculations. SPT = Site productivity taxation, RIT = realized income taxation, N = the average number of the landowners in seven different models, Stdev. = the standard deviation of the number of landowners in different models and % = the average percentage of the correctly modelled choice.

Logistic regression				Neural networks			
Modelled / actual	N	Stdev.	%	Modelled / actual	N	Stdev.	%
SPT / SPT [1)]	21	2.06	**87**	**SPT / SPT** [1)]	16	5.82	**66**
RIT / SPT	3	1.57		RIT / SPT	8	5.87	
SPT / RIT	15	3.95		SPT / RIT	14	10.58	
RIT / RIT	19	3.69	**57**	**RIT / RIT**	20	10.39	**60**
Total	58		**69**		58		**62**

[1)] E.g., the modelled choice was site-productivity taxation and the actual choice was site-productivity taxation.

In the modelling the two choices, the standard deviations of the numbers of landowners in different classified groups were higher in modelling with neural networks than with logistic regression. Moreover, the standard deviations were higher in modelling the second choice than the first choice. The standard deviations were also higher in estimating the choice of RIT than of SPT (Tables 3 and 4).

The analysis of the classifications of logistic regression and neural networks was done using t-tests. In the analyses, the difference between the numbers of landowners in each class was tested between the two methods (e.g. for the first choice, both the actual and predicted choices were SPT in both methods). Differences at 5 % significance level were observed between the classes with the second choice being SPT (Table 4).

5 Discussion

5.1 Predicting the Choice of Forest Taxation Basis

In this study, in the comparison between logistic regression and neural networks, logistic regression produced better results in modelling both choices. On average, logistic regression predicted nearly 70 % of the choices correctly, while neural networks fared almost 10 percentage units worse. In all the models applied, the choice of site-productivity taxation was predicted more correctly than the choice of realized-income taxation. This can interpreted in two ways: either the landowners were more certain about their choice of SPT, or they intended to cut timber considerably more than the level of cuttings where the profitability of the forest taxation basis changes.

Cross-validation was used to increase the reliability of the results obtained. In modelling, both data sets were randomly divided into seven training sets and seven test sets. The standard deviations of the correctly and incorrectly estimated choices for neural networks were higher than for logistic regression equations. Thus, compared to logistic regression, neural networks are more sensitive to the division of data into training and testing sets. Had the data set been divided only once, the results for neural networks would have varied markedly with a relatively small number of observations. The reliability of the results provided by neural networks could be improved by increasing the number of observations in the analyses.

According to the χ^2 values obtained, the logistic regression model performed better in modelling the first choice than the second choice. It can be assumed that landowners with a strong opinion on the preferred, but not always the optimal, forest taxation basis in the first inquiry did not change their opinion in the second inquiry – regardless of what seemed to be the optimal taxation system. Landowners with no opinion on the preferred forest taxation basis in the first inquiry were divide equally between the choices of optimal and non-optimal forest taxation bases in the second inquiry.

In analysing the log-odds of the independent variables in modelling the first choice, the probability of choosing SPT increased when the cuttings during the transition period increased. Young landowners, intent on harvesting in the near future, chose SPT more frequently than RIT. Furthermore, in modelling the second choice, the probability of choosing SPT increased when the planned harvest rates increased during the transition period. Increase in the marginal tax rate of the landowners decreased the probability of choices favouring SPT. The results are consistent with, for example, Pesonen and Räsänen (1993).

In modelling the first choices (before calculations), the factors affecting the choice of taxation basis were:

1) forest taxation index,
2) age of landowner, and
3) intention to cut timber during the next three years

Respectively, the significant factors in modelling the second choice (after calculations) were:

1) forest taxation index,
2) estimated personal marginal tax rate in 1991, and
3) intention to cut timber during the next three years.

The most important variable affecting the choice of forest taxation basis was the forest taxation index. The connection between timber management strategies and the choice of forest taxation basis has been reported by Pesonen (1995) and Pesonen et al. (1995). The choice of forest taxation basis is part of the strategic decision making of NIPF landowners. According to the results, most landowners were aware of the most important factor affecting the optimal choice, i.e. first they chose the preferred timber management strategy, and after that, the suitable forest taxation basis. Also, the connection between intention to cut during the next three years and the choice of the forest taxation basis was apparent, i.e. for landowners with no intention to cut during 1993–1995, the probable choice was RIT.

In general, the characteristics of landowners did not demonstrate any marked connection with the choice of forest taxation basis. The landowner's age was proven to be slightly significant in modelling the choices before any information was produced about the economic consequences of alternative taxation bases for the transition period. It is inevitable that during the life cycle of NIPF landowners, the strategic goals will vary. Furthermore, Kuuluvainen and Salo (1991) have reported that age is a significant variable when analysing the timber supply and life cycle harvest of NIPF landowners.

The intention to cut timber during the years 1993–1995, the estimated personal marginal rate in 1991, and the forest taxation index, all of which affected the optimal choice of the forest taxation basis, were more significant in the modelling of the second choices, i.e. after NIPF landowners had been shown information about the economic consequences of their choices. In the modelling of the second choice, the landowner's marginal tax rate was more significant than in the models for the first choice. In addition, the non-economical factors lost their significance in modelling the second choice. In that sense, the calculations increased the economic consciousness of the landowners.

The material used in this study was somewhat limited in terms of the point in time when the inquiry was conducted. At the time as the material was collected, the summer of 1993, the final decisions of concerning the forest taxation basis for NIPF landowners were not known. In addition, the final data sets used were quite small, consisting of only 147 and 193 NIPF landowners respectively. The accuracy of the models could have been improved with the use of final and actual choices and larger sample sizes.

5.2 Conclusions

In this study, logistic regression predicted the choice of forest taxation basis more accurately than did neural networks. Logistic regression gives good possibilities for quantifying independent variables and their influence on the dependent variable. In general, the modelling of NIPF landowners' behaviour is considered to be a multi-dimensional problem with few possibilities for developing accurate models (e.g. Pesonen et. al. 1995).

In further studies, it is important to study landowners' behaviour during the transition period. There are numerous possibilities for comparing landowners who chose site-productivity taxation with those who chose realized-income taxation. Furthermore, other interesting issues for study would be the possible differences in the future timber cutting behaviour of landowners, and the effects that the Finnish forest taxation reform has on the future supply of timber from NIPF landowners. Moreover, an essential matter to find out is the division of potential, allowable cut between landowners differentiated by their choice of forest taxation basis.

After obtaining the actual choices of forest taxation basis, the possible differences in the distribution of the choices in different parts of Finland could be clarified. If the choices vary, the crucial subject to study would then be the reasons for this variation.

In the future, it will be possible to test other machine learning techniques, e.g. genetic algorithms, in modelling the choice of forest taxation basis. As an example of alternative machine learning approaches, Pesonen et al. (1995) modelled the NIPF landowners' choices of timber management strategies using a genetic algorithm.

Acknowledgements

This study is part of the project 'Optimization of Regional Cutting Budgets' in The Finnish Forest Research Institute. The main objective of the project is to develop a new system of calculation for determining regional cutting budgets. Special thanks are due to the following partners of the project for their help and advice during the research process: The Finnish Forest Industries Federation, The Central Union of Agricultural Producers and Forest Owners, The Forest Center Tapio, The Forestry District of Pohjois-Savo, The University of Joensuu and The National Board of Taxation.

The authors would also like to thank the following persons for their helpful comments on the manuscript: Veli-Pekka Heikkinen, M.Sc.(Econ.), Hannu Hirvelä, M.Sc.(For.), Arto Kettunen, M.Sc.(For.), Dr. Ville Ovaskainen, Dr. Timo Pukkala, and Olli Salminen, M.Sc.(For.). Thanks are also extended to Mr. E. Pekkinen for checking the English language.

References

Bailey, D. & Thompson, D. 1990. How to develop Neural Network. AI Expert, June 1990: 9–47.
BMDP statistical software. 1992. BMDP Statistical Software, Department of Biomathematics, UCLA, Los Angeles. University of California Press, Berkeley–Los Angeles–London.
Cook, D. & Wolfe, M. 1991. A back-propagation neutral network to predict average air temperatures. AI Applications 5(1): 40–46.
— , Massey, J.G. & Shannon, R. 1991. A neutral network to predict particleboard manufacturing process parameters. Forest Science 37(5): 1463–1478.
Demaris, A. 1992. Logit modeling: practical applications. Sage University Paper Series on Quantitative Applications in the Social Sciences, 07-086. Sage, Newbury Park, CA.
Feychting, H., Hagner, O. & Nilsson, M. 1991. Estimation of forest stand characteristics using neural networks and satellite remote sensing. In: Saarenmaa, H. (ed.). Current advances in the use of computers in forest research. Metsäntutkimuslaitoksen tiedonantoja 395: 13–17.
Greene, W. 1993. Economic analysis. Macmillan Publishing Company, New York. 791 p.
Guan, B. & Gertner, G. 1991. Machine learning and its possible roles in forest science. AI Applications, Natural Resources, Agriculture, and Environmental Science 5(2): 27–36.
Heliövaara, K, Väisänen, R. & Immonen, A. 1991. Quantitative biogeography of the bark beetles (Coleoptera, Scolytidae) in Northern Europa. Acta Forestalia Fennica 219. 35 p.
Johnson, K.N., Stuart, T.W. & Crim, S.A. 1986. FORPLAN version II: an overview. Land Management Planning Systems Section, USDA Forest Service, Washington, DC.
Jonsson, B., Jacobsson, J. & Kallur, H. 1993. The forest management planning package. Theory and application. Studia Forestalia Suecica 189. 56 p.
Kajanus, M. 1992. Metsänuudistamispäätökset osana metsätalousyrityksen johtamista. Licentiate thesis. University of Helsinki. 71 p.
Kangas, J. & Pukkala, T. 1992. A decision theoretic approach applied to goal programming of forest management. Silva Fennica 26: 169–176.
Kast, E.K. & Rosenzweig, J.E. 1974. Organization and management. A systems approach. Second edition. 389 p.
Kilkki, P. 1968. Income-oriented cutting budget. Acta Forestalia Fennica 91. 54 p.
— & Siitonen, M. 1976. Principles of a forest information system. XVI IUFRO World Congress, Division IV, Proceedings. p. 154–163.
Kuuluvainen, J. & Salo, J. 1991. Timber supply and life cycle harvest of nonindustrial private forest owners: an empirical analysis of the Finnish case. Forest Science 37(4): 1011–1029.
Laki maatilatalouden tuloverolain muuttamisesta. 1990. Suomen säädöskokoelma n:o 718.
Lämås, T., Ståhl, G. & Thuresson, T. 1991. Estimating inoptimality losses in harvesting decisions using neural networks. In: Saarenmaa, H. (ed.). Current advances in the use of computers in forest research. Metsäntutkimuslaitoksen tiedonantoja 395: 5–12.
Lönnstedt, L. 1989. Goals and cutting decisions of private small forest owners. Scandinavian Journal of Forest Research 4: 259–265.
— & Törnqvist, T. 1990. The owner, the estate and the external influences: nonindustrial private forest owners' management decisions. Swedish University of Agricultural Sciences, Department of Forest-Industry-Market Studies, Report 14.
— & Roos, A. 1993. Cutting levels for nonindustrial forestry: an estimate based on the state of forests and the goals of forest owners. Swedish University of Agricultural Sciences, Department of Forest-Industry-Market Studies, Report 26.
Lindroos, H. 1992. Metsäverouudistus ja puuhuolto. In: Pesonen, M. & Räsänen, P. (eds.). Metsäverovalinta – strateginen ratkaisu. Metsäntutkimuslaitoksen tiedonantoja 472.
NeuroShell. 1989. Neural Network Shell Program. Manual. Ward Systems Group, Inc.
Nikula, A. & Väkevä, J. 1991. Modelling the risk of moose browsing in forest plantations with neural networks. In: Saarenmaa H. (ed.). Current advances in the use of computers in forest research. Metsäntutkimuslaitoksen tiedonantoja 395: 18–21.
Ovaskainen, V., Hänninen, H. & Kuuluvainen, J. 1992. Puunmyyntitulojen verotukseen siirtymisen vaikutukset puun tarjontaan ja verokertymään. Metsäntutkimuslaitoksen tiedonantoja 440. 23 p.
Pääomatulojen verotuksen ja yritysverotuksen kehittämislinjat. 1991. Asiantuntijatyöryhmän muistio. Valtiovarainministeriön työryhmämuistioita 1991: 28. 116 p.

Pattie, D. 1992. Using neural networks to forecast recreation in wilderness areas. AI Applications 6(2): 57–59.
Pesonen, M. 1995. Non-industrial private forest landowner's choices of timber management strategies and potential allowable cut: case of Pohjois-Savo. Acta Forestalia Fennica 247. 31 p.
— , Kettunen A. & Räsänen P. 1995. Non-industrial private forest landowner's choice of timber management strategy: genetic algorithm in predicting harvest rates. Manuscript. Finnish Forest Research Institute.
— & Räsänen, P. (eds.). 1993. Metsäverovalinta – strateginen ratkaisu. Metsäntutkimuslaitoksen tiedonantoja 472.
Pukkala, T. & Kangas, J. 1993. A heuristic optimization method for forest planning and decision making. Scandinavian Journal of Forest Research 8: 560–570.
Ranta, R. 1991. Metsätalouden suunnittelulaskelmat. In: Tapion taskukirja. 21st edition. Kustannusosakeyhtiö Metsälehti. p. 334–337.
Royer, J. 1987. Determinants of reforestation behavior among southern landowners. Forest Science 33(3): 654–667.
Rumelhart, D., Hinton, G. & McClelland, J. 1986. Learning internal representations by error propagation. In: Rumelhart, D.E., McClelland, J.L. & the PDP Research Group (eds.). Parallel distributed processing. Explorations in the microstructure of cognition. Volume 1: Foundations. MIT Press, Cambridge, Massachusetts–London, England.
Saaty, T.L. 1977. A scaling method for priorities in hierarchical structures. Journal of Mathematical Psychology 15: 234–281.
— 1980. The analytic hierarchy process. McGraw-Hill. New York. 287 p.
Siitonen, M. 1983. A long term forestry planning system based on data from Finnish National Forest Inventory. In: Forest inventory for improved management. Helsingin yliopiston metsänarvioimistieteen laitoksen tiedonantoja 17: 195–207.
— 1993. Experiences in the use of forest management planning models. Silva Fennica 27(2): 167–178.
Tuloverolaki. 1992. Suomen säädöskokoelma n:o 1535.
Uusitalo, J. 1993. Sahatavaran laadun ennustaminen mäntytukkirungosta. Licentiate thesis. University of Helsinki. 63 p.
Wardle, P.A. 1965. Forest management and operational research: a linear programming study. Ser. B, Management Science 11(10): 260–270.
Ware, G.O. & Clutter, J.L. 1971. Mathematical programming system for the management of industrial forests. Forest Science 17(4): 428–445.

Total of 43 references

Appendix. Logistic models for the choices of forest taxation basis before and after calculations with secen independent variables.

Variable	Before calculations		After calculations	
	Beta	Coeff./SE	Beta	Coeff./SE
Intercept	–2.6360	–1.67	–1.0930	–0.98
Age	–0.0233*	–1.41	–0.0169	–0.86
Farmer	0.2655	0.58	0.1287	0.363
Area	–0.0019	–0.58	–0.0001	–0.02
Cuttings	–0.0972	–0.94	–0.0308	–0.40
Cu9395	1.9520**	1.80	1.2380***	2.11
Rate	–0.0126	0.81	–0.0293***	–2.48
Index	1.1420***	3.87	0.8242***	3.70
χ^2 [1]	6.047		14.621	
Df	8		8	
N	147		193	

*** denotes significance at 0.05 level
** denotes significance at 0.10 level
* denotes significance at 0.20 level

[1] The test is Hosmer-Lemeshow test because some of the variables are continuous. The Hosmer-Lemeshow goodness of fit chi-square test divides the data into 10 cells and compares the observed and predicted frequencies for these cells. The cells are defined using the predicted frequences (DeMaris 1992, BMDP... 1992).

Submission of Manuscripts Manuscripts should be submitted in triplicate to Silva Fennica, Unioninkatu 40 A, FIN-00170 Helsinki, Finland. Detailed instructions to authors are printed in the first issue each year, and can be found on WorldWideWeb at URL http://www.metla.fi/publish/silva/SF-Instructions.html/. Offprints of Instructions are available on request.

Publication Schedule Silva Fennica is issued in four numbers per volume.

Subscriptions and Exchange Subscriptions and orders for back issues should be addressed to Academic Bookstore, Subscription Services, P.O. Box 23, FIN-00371 Helsinki, Finland, Phone +358 0 121 4430, Fax +358 0 121 4450. Subscription price for 1995 is 300 FIM (for subscribers in Finland 200 FIM). Exchange inquiries should be addressed to The Finnish Society of Forest Science, Unioninkatu 40 B, FIN-00170 Helsinki, Finland, Phone +358 0 658 707, Fax +358 0 191 7619, E-mail sms@helsinki.fi

Statement of Publishers Silva Fennica has been published since 1926 by The Finnish Society of Forest Science. From 1994, the journal is published by the Finnish Society of Forest Science and the Finnish Forest Research Institute. The Finnish Society of Forest Science is a nonprofit organization founded in 1909 to promote forest research. The Finnish Forest Research Institute, founded in 1917, is a research organization financed by the Ministry of Agriculture and Forestry.

Abstracting Articles in Silva Fennica are abstracted and indexed in Agrindex, Biological Abstracts, Current Advances in Ecological Sciences, Current Advances in Plant Sciences, Ecological Abstracts, Forest Products Journal, Forestry Abstracts, International Bibliography of Periodical Literature, Life Sciences Collection.

SILVA FENNICA Vol. 29(2), 1995

The Finnish Society of Forest Science
The Finnish Forest Research Institute

Research articles

ISSN 0037-5330

9 770037 533017