Agronomy Research **3**(2), 177–187, 2005

Artificial neural network as an alternative to multiple regression analysis for estimating the parameters of econometric models

R. Põldaru[1], J. Roots[1] and A.-H. Viira[2]

[1]Institute of Economics and Social Sciences,
Estonian University of Life Sciences, Kreutzwaldi 64, 51014 Tartu, Estonia;
e-mail: Reet.Poldaru@emu.ee, Jyri.Roots@emu.ee
[2]ARIB, Narva 3, 51009, Tartu, Estonia; e-mail: Ants.Viira@pria.ee

Abstract. In recent years, neural networks have been used for a wide variety of applications where statistical methods are traditionally employed. Neural nets offer the opportunity to create a model by using technology similar to the learning patterns of the human brain. The structure of artificial neural networks (ANN) is based on the human brain's biological neural processes. Artificial neural networks provide a new approach to the problem of parameter estimation of nonlinear econometric models. This paper presents a comparison between neural networks and econometric approaches for estimation of parameters of an econometric model of grain yield. The aim of this study is to show that neural nets are a convenient econometric tool. The parameters were estimated on the basis of alternative variants of models. The analysis shows that artificial neural network models may be used for parameter estimation of the econometric models.

Key words: artificial neural network, econometric models, grain yield

INTRODUCTION

An artificial neural network is a learning system based on a computational technique, which attempts to simulate the neurological processing ability of the brain. Initially, neural networks were developed as simulation models of brains. The terminology used is still a reminder of this origin. ANN could be applied to quantify a non-linear relationship between casual factors.

Artificial neural networks are a class of models developed by cognitive scientists interested in understanding how computation is performed by the brain. They nevertheless provide a rich, powerful and interesting modelling framework with proven and potential application across the sciences. Neural networks are used in many sciences like biology, informatics, econom(etr)ics (Kuan & White, 1994; Kaashoek & van Dijk, 2000; La Rocca & Perna, 2005), and agriculture (Kominakis et al., 2002; Põldaru & Roots, 2003; Yang et al. 2003; Kaul et al., 2005; Park et al., 2005; Uno et al., 2005).

In this paper, we concentrate on the possibility of implementing artificial neural networks for the estimation of econometric model parameters of grain yield.

In the current investigation, we used multiple linear regression method and an ANN to estimate the parameters of econometric model of grain yield.

The aim of this study is the estimation of the parameters of an econometric model of grain yield and the analysis of results. The paper provides an overview of artificial neural networks, describes their potential implementation in rural areas, and discusses the implementation of this method for analysing the field crops sector in Estonia.

MATERIALS AND METHODS

The data used are an unbalanced panel of grain growers drawn from the FADN (Farm Accountancy Data Network) database of Estonian grain growers. The parameters are estimated on the basis of alternative models of artificial neural network by using a non-linear model. The results are compared mutually and with results of multiple regressions.

The dependent variable is average grain yield (y), and independent variables are time dummies (x_1, x_2, x_3 and x_4); variables for input costs: seeds (x_5), fertilisers (x_6), crop protection (x_7) and other crop specific costs (x_8); variable of machinery cost (x_9), variable of current assets (x_{10}), variable of land quality (x_{11}), and variable of production structure (x_{12}) (Table 1).

Table 1 provides the coefficients of correlation between inputs (independent variables) and output (grain yield). The descriptive statistics of data used for model parameter estimation is reported in Table 2.

For estimating the parameters of the econometric model of grain yield, the Excel Solver and data analysis software system STATISTICA version 7 are used.

In this study, ordinary least squares (multiple regression method) and artificial neural network models for grain yield model parameter estimations are compared.

Artificial neural networks

The basic elements of a neural network comprise neurons and their connection strengths (weights). Neurons are grouped into layers (see Figure 1). In a multi-layer network there are usually an input layer, one or more hidden layers, and an output layer. Between the input and output layer one can have a hidden layer that is used to solve nonlinear problems. The network is fully connected. That is all the nodes are linked in adjacent layers. These links are the weights that can be strong or weak, depending on their value. The weights are adjusted for the minimisation of the mean square error as the objective function.

The cells of the input layer correspond to the “regressors” or “independent variables” in the standard linear regression model. The cells in the output layer correspond to the dependent variables in the linear model. The hidden layer contains cells, which transmit the signals from the input layer to the output layer. These cells may be interpreted as unobserved components built into the linear model. It is the presence of this hidden layer that permits the nonlinear mapping since similar networks lacking a hidden layer can only affect a multivariate linear mapping (Kaashoek & Van Dijk, 2000).

A graph of a neural network with five cells in the input layer, two cells in the hidden layer and one cell in the output layer is shown in Figure 1.

The network transmits the signals as follows: A weighted sum of the signals of the input cells is sent to the hidden layer cells. Within the cells of this layer, the values of the signals received are transformed by the so-called 'activation function'. Then a weighted sum of the transformed signals is sent to the cells of the output layer. The weights in the neural network correspond to unknown parameters in the linear model.

Table 1. Definitions of independent variables (inputs).

	Definitions of independent variables (inputs)	Unit	Xi	Coefficient of correlation
1	Dummy variable for year 1999	-	x_1	-0.28
2	Dummy variable for year 2000	-	x_2	0.10
3	Dummy variable for year 2001	-	x_3	0.04
4	Dummy variable for year 2002	-	x_4	0.09
5	Seed	EEK ha^{-1}	x_5	0.18
6	Fertilisers	EEK ha^{-1}	x_6	0.45
7	Crop protection	EEK ha^{-1}	x_7	0.32
8	Other crop specific costs	EEK ha^{-1}	x_8	0.22
9	Machinery cost	EEK ha^{-1}	x_9	0.17
10	Current assets	EEK ha^{-1}	x_{10}	0.28
11	Quality of land	points	x_{11}	0.13
12	Fraction of grain sown area in total sown area	%	x_{12}	-0.04

Table 2. Descriptive statistics of information (independent variables) used for this study.

Independent variables (inputs)	Mean	Minimum	Maximum	Median	SD
Seed	505	80	1,585	433	256
Fertilisers	913	607	1,729	888	217
Crop protection	394	0	1,438	382	212
Other crop specific costs	128	0	3,803	36	312
Machinery cost	2,772	0	21,847	2,098	2,823
Current assets	2,405	0	9,398	1,979	1,716
Quality of land	43	37	50	45	4
Fraction of grain sown area in total sown area	505	80	1,585	433	256

Table 3. Model coefficients estimated by multiple linear regression.

Xi	Factor	Coefficient	Computed *t*-value	*P*-value
x_0	Intercept	195.6345	0.394771	0.693389
x_1	Dummy variable for year 1999	-540.08	-3.05277	0.002543*
x_2	Dummy variable for year 2000	468.2803	3.711402	0.00026*
x_3	Dummy variable for year 2001	181.9562	1.628458	0.10484
x_4	Dummy variable for year 2002	220.2581	2.072925	0.039328*
x_5	Seed	-0.12147	-0.74551	0.456749
x_6	Fertilisers	0.976527	4.681037	4.95E-06*
x_7	Crop protection	0.679227	3.398298	0.000803*
x_8	Other crop specific costs	0.121439	0.937605	0.349462
x_9	Machinery cost	-0.00601	-0.37392	0.70882
x_{10}	Current assets	0.115587	4.36595	1.94E-05*
x_{11}	Quality of land	17.11607	1.725807	0.085767
x_{12}	Fraction of grain sown area in total sown area	0.454819	0.209284	0.834417

* *Significant at P* < 0.05

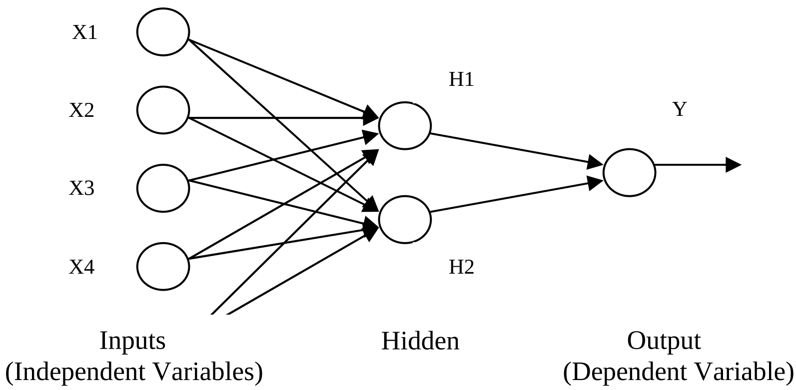

Fig. 1. Layers and connections of an artificial neural network.

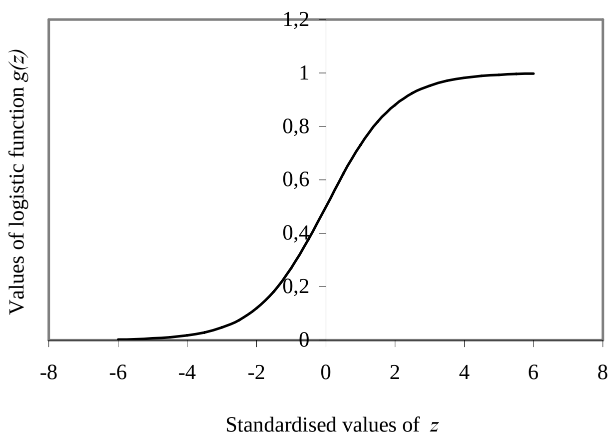

Fig 2. Graph of logistic activation function.

Table 4. Characteristics of different neural network models.

Model identificator	Model type	Number of nodes in input layer	Number of inputs to hidden layer	Number of nodes in hidden layer	Total number of parameters
ANN1	1	11	11	1	14
ANN1L	2	11	7	1	14
ANN2L	2	11	7	2	23

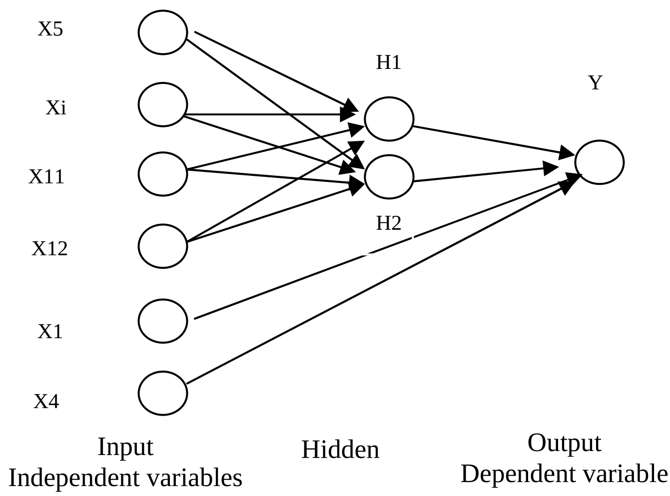

Fig. 3. Graph of a neural network for second model.

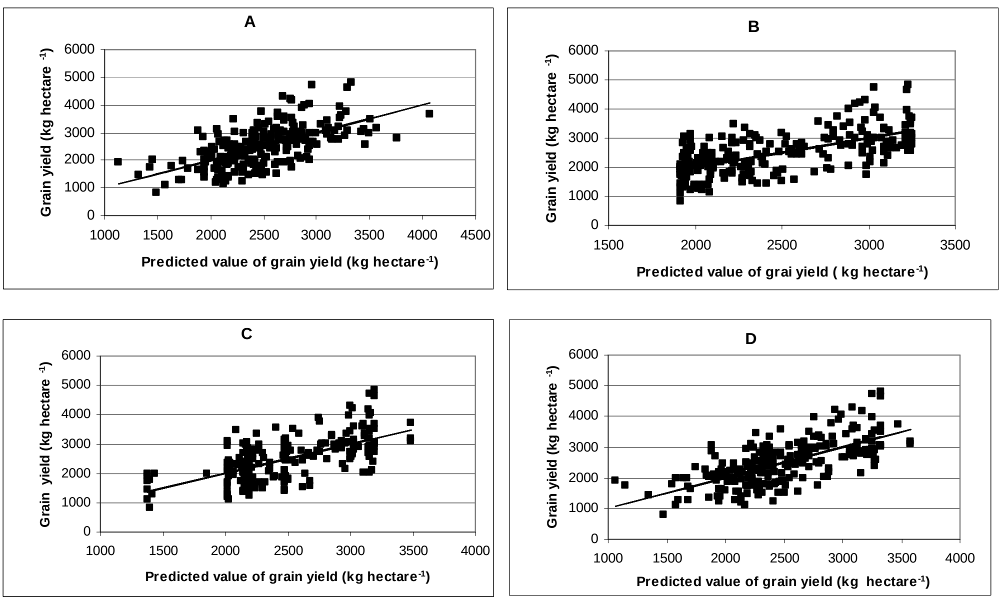

Fig 4. Relationships within grain yield and predicted grain yield for different model variants: OLS (A), ANN1 (B), ANN1L (C), and ANN2L (D).

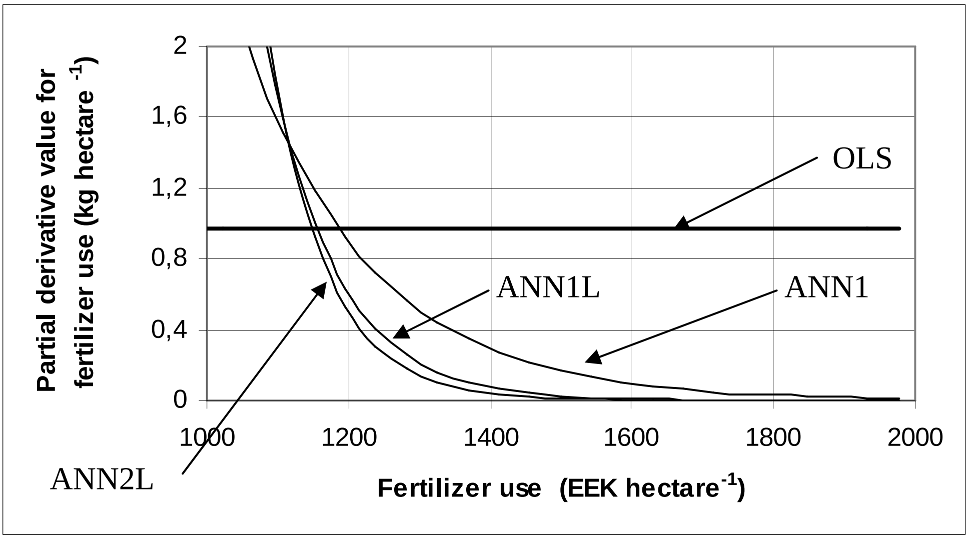

Fig 5. Graph of derivatives with respect to values of independent variable for fertiliser use.

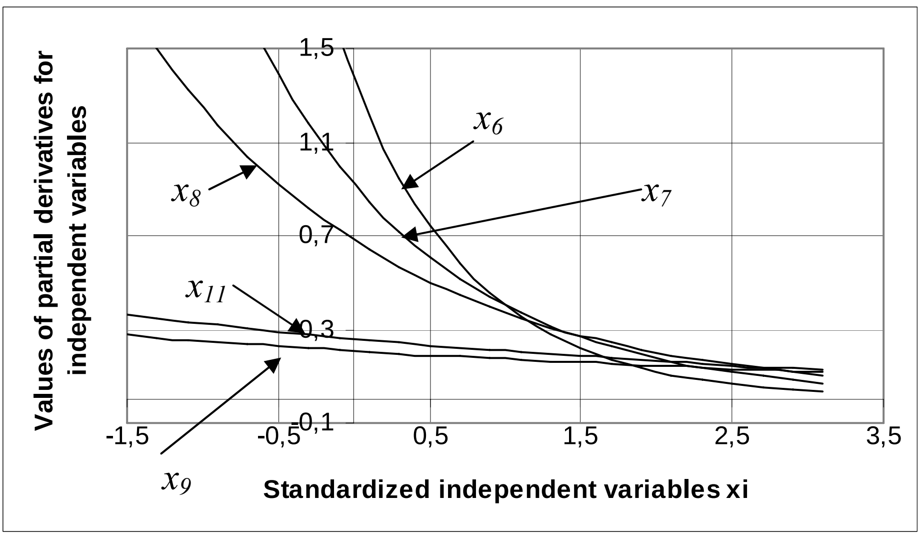

Fig. 6. Graph of derivatives with respect to values for different independent variables.

Table 5. Summary characteristics of different models.

	OLS	ANN1	ANN1L	ANN2L
Observations	236	236	236	236
Number of parameter	12	14	14	23
Standard error (residuals)	580.4	564.6	564.8	554.5
R square	0.382	0.418	0.417	0.461
Multiple R	0.618	0.646	0.646	0.679
Adjusted R square	0.349	0.384	0.383	0.405
F	11.48	12.04	12.02	8.74
P-value	0.000	0.000	0.000	0.000

In this study, two different ANN models are investigated. The mathematical structure of a simple neural net may write out as follows. The value *y* is equal to

$$y = \sum_{h=1}^{H} c_h \cdot g\left(\sum_{i=1}^{I} a_{ih} \cdot x_i + b_h \right) + d \qquad (1)$$

where

- i index of input cells (independent variables), $i = 1;...; I$;
- h index of hidden layer cells, $h = 1;...;H$;
- $g(z)$ activation function;
- x_i value of input cell (explanatory variables) i;
- y value of output cell (dependent variable);
- a_{ih} weight of the signal from input cell i to hidden cell h;
- b_h constant input weight for hidden cell h;
- c_h weight of the signal from hidden cell h to output cell y;
- d constant weight for output cell.

Generally, a nonlinear sigmoidal/logistic function is used to regulate the output of a node, shown as follows:

$$g(z) = \frac{1}{1+\exp(-z)} \qquad (2)$$

where $g(z)$ is the output of the node in the hidden layer.

The sigmoid function producing outputs in the range of [0,1] is used as a transfer characteristic for each neuron in the hidden and output layers. Figure 2 is a representation of the graph of the function (2) (Dunham, 2003).

For the second model, the value y of the output cell is given as the sum of weighted value of the output of the hidden cells and the sum of weighted value of dummy variables. It is equal to

$$y = \sum_{h=1}^{H} c_h \cdot g\left(\sum_{i=1}^{I_1} a_{ih} x_i + b_h \right) + \sum_{j=1}^{I_2} a_j x_j + d \qquad (3)$$

where the activation function $g(z)$ is logistic function (2) and I_2 is the number of input variables. For our model $I_2 = 4$. The principal difference is that dummy variables (x_1, x_2, x_3 *and* x_4) are not activated by hidden cell. The other variables ($x_5 \ldots x_{12}$) are activated by different numbers of cells in the hidden layer. In the second model (3) the first part is nonlinear and the second part is linear.

For estimation of parameters of neural networks, a generally accepted optimisation principle is to minimise the mean squared prediction error,

$$E_D = \frac{1}{N} \sum_{t=1}^{N} \left(y_t - \bar{y}_t \right)^2 \qquad (4)$$

where t is index of output y values (sample observations) $t = 1;\ldots;N$.

Hereafter, for brevity, we denote the model parameters c_h, a_{ih}, b_h and d by w_i. To find these parameters, you have to solve a quadratic programming problem. The Excel Solver may be used to find the best model parameter values. The control parameters α and β can be used to stabilise the optimisation process (MacKay, 1995).

The modified objective function is equal to

$$M(w) = \beta \cdot E_D + \alpha \cdot E_w , \qquad (5)$$

whereas $$E_w = \frac{1}{2} \cdot \sum_{i=1}^{W} w_i^2 \qquad (6)$$

The use of the sum-squared error E_D as defined above (4) corresponds to the assumptions of Gaussian noise on the dependent variable, and the parameter β defines a noise level. If E_w is quadric as defined above (6), then the potential fluctuations in parameter estimates can be decreased and the optimisation process regularised.

In this paper, two different (alternative) models were investigated. The first model was the classical neural network model (Figure 1 and formulas (1) and (2)), where the number of cells in the hidden layer was different (1...3).

The graph of the second model is shown in Figure 3. The principal difference is that the dummy variables (x_1, x_2, x_3 *and* x_4) are not activated by the hidden cell. The other variables ($x_5 \ldots x_{12}$) were activated by different numbers of cells in the hidden layer.

Neural networks provide a new approach to the problem of parameter estimation of nonlinear models.

RESULTS AND DISCUSSION

For the modeling of grain yield, three different ANN models are used. The summary of the characteristics of the models is reported in Table 3. Table 3 shows the model parameters including the number of nodes in the hidden layer and the total number of parameters.

For assessing the ANN models, the results are compared with the ordinary least squares (OLS) model. Table 4 presents the estimates of parameters of the econometric model estimated on the basis of OLS.

The values in Table 3 indicate that

- The linear model gives acceptable results (the signs of parameter estimates are in accordance with economic theory)
- Most of the parameter estimates are statistically very reliable (t – statistics are very high)

Table 5 presents summaries of the results of various model alternatives of ANN. Summary characteristics for various alternatives are standard error, the coefficient of determination R^2, the adjusted coefficient of determination R^2, and F-criterion.

Let us discuss the summary characteristics in Table 5. The number of parameters is different for different alternatives. The prediction accuracy of the models (values of standard error) is practically the same. The values of the coefficient of determination R^2 are relatively high. Table 5 shows that the value of determination coefficients depends on the model type. Minimum value is for OLS model (0.382) and maximum value for ANN2L model (0.461). Hereby, all ANN models have higher value than OLS models. All models are statistically significant. Consequently, the ANN models work well.

Figure 4 illustrates the scatter plots of grain yield depending on predicted values for different model alternatives. The continuous line on the graphs is the predicted value of grain. The distance of points from the continuous line is a residual. Figure 4 shows that the graphs are different for different models. Relatively analogous are the graphs for OLS model (A) and ANN2L model (D). Most different is the graph for ANN1 model (B). For alternatives (A) and (D) the graphs have a typical shape, the intensity of the points diminishes in both ends of the graph, but for ANN1 model the intensity of the points is relatively high at the ends of the graph. In that case all independent variables are activated by using the sigmoid function.

Next we calculate and discuss the values of partial derivatives with respect to independent variable of fertiliser use for different variants of the econometric model. The values of partial derivatives are computed from predicted values numerically.

The graph of derivatives with respect to independent variable for fertiliser use is shown in Figure 5. The dot line on the graphs presents an OLS regression coefficient (see Table 2).

Figure 5 indicates that the values of partial derivative differ for different models considerably. For example, in the case of OLS model, the values of partial derivative do not depend on values of independent variables. For all variants, the values of independent variables are the same – 0.97. Consequently, the linear model is the most inflexible one.

In the case of neural network models, the values of partial derivative depend on nonlinearity of the values of fertiliser use. The graphs for different ANN model alternatives do not differ substantially. When the independent variable has a low value then the partial derivative has a higher value. When the independent variable has a high value then the partial derivative has a lower value. This regularity is in accordance with economic theory. The economic theory asserts that with increasing the use of the resource, the effectiveness of the resource must diminish.

Next we analyse the partial derivative of grain yield with respect to most influential independent variables (fertiliser use - x_6, crop protection – x_7, other crop specific costs – x_8, machinery costs – x_9, and quality of land – x_{11}) for ANN1 model.

The graph of derivatives with respect to independent variables is shown in Figure 6. Since the independent and dependent variables are measured in different dimensions, the independent and dependent variables on the graph are presented in standardised values (units).

Looking at the graph, we may now summarise the following conclusions:

- Some relations are essentially nonlinear. For example, the partial derivative of grain yield with respect to independent variable x_6 (fertiliser use), x_7 (crop protection) and x_8 (other crop specific costs) have the essential nonlinearity (curvature). The graphs of other partial derivatives are almost linear.
- All the graphs are declining. From economic point of view, the graph of derivative of econometric function with respect to independent variables of economic resources (land quality, fertiliser use, etc) must be decreasing. Consequently, the ANN models are in accordance with economic theory.

CONCLUSIONS

Neural network is a fast expanding field with many new research results reported and developed recently. The results show that neural network models are flexible and informative. The ANN models gave acceptable results and may be recommended for practical use. The neural network models may be used for estimating the parameters of nonlinear econometric models and implemented in agricultural research.

REFERENCES

Dunham, M. H. 2003. *Data Mining Introductory and Advanced Topics.* Pearson Education, Indian Branch, India, 315 pp.

Kaashoek, J.F. & van Dijk, H. K. 2000. Neural networks as econometric tool. *Econometric Institute Rapport EI2000-31A.* Econometric Institute, Erasmus University Rotterdam. Rotterdam, 29 pp.

Kaul, M., Hill, R.L. & Walthall, C. 2005. Artificial neural networks for corn and soybean yield prediction. *Agricultural Systems* **85**(1), 1–18.

Kominakis, A.P., Abas, Z., Maltaris, I. & Rogdakis, E. 2002. A preliminary study of the application of artificial neural networks to prediction of milk yield in dairy sheep. *Computers and Electronics in Agriculture* **35**(1), 35–48.

Kuan, C.-M. & White, H. 1994. Artificial Neural Networks: An Econometric Perspective. *Econometric Reviews* **13**, 1–92.

La Rocca, M. & Perna, C. 2005. Variable selection in neural network regression models with dependent data: a subsampling approach. *Computational Statistics & Data Analysis* **48**(2), 415–429.

MacKay, D.J.C. 1995. Bayesian Methods for Neural Networks: Theory and Applications. *Course notes for Neural Networks Summer School,* 43 pp. http://www.inference.phy.cam.ac.uk/mackay/cpi4.pdf.

Park, S.J., Hwang, C.S. & Vlek, P.L.G. 2004. Comparison of adaptive techniques to predict crop yield response under varying soil and land management conditions. *Agricultural Systems* **85**(1), 59–81.

Põldaru, R. & Roots, J. 2003. The Estimation of the Econometric Model of Grain Yield in Estonian Counties Using Neural Networks. *VAGOS* **57**(10). Akademija, Kaunas, pp. 124–130.

Uno, Y., Prasher, S.O., Lacroix, R., Goel, P.K., Karimi, Y., Viau, A. & Patel, R.M. 2005. Artificial neural networks to predict corn yield from Compact Airborne Spectrographic Imager data. *Computers and Electronics in Agriculture* **47**(2), 149–161.

Yang, C.-C., Prasher, S.O, Landry, J.-A. & Ramaswamy, H.S. 2003. Development of a herbicide application map using artificial neural networks and fuzzy logic. *Agricultural Systems* **76**(2), 561–574.