Forecasting of cut Christmas trees with Artificial Neural Networks (ANN)

Ioannou Konstantinos[1], Garyfallos Arabatzis[2], Theodoros Koutroumanidis[3] and Georgios Apostolidis[2]

[1]Faculty of Forestry and Natural Environment, Aristotle University of Thessaloniki, e-mail: ioannou.konstantinos@gmail.com
[2]Department of Forestry and Management of the Environment and Natural Resources, Pantazidou 193, Orerstiada, e-mail: garamp@fmenr.duth.gr
[3]Department of Agricultural Development, Pantazidou 193, Orerstiada, e-mail: tkoutrou@agro.duth.gr

Abstract. The establishment of Christmas tree plantations over the last 50 years in Greece has provided an additional income for the inhabitants of certain mountainous and semi-mountainous regions of the country and contributed to their development. In this paper is used a neural network which combined with ARIMA which makes forecast of number of actually cut Christmas trees until the year 2015.

Keywords: TNN, ARIMA, Christmas Trees, Greece

1. Introduction

Forestry is very closely linked to the economy of mountainous and semi-mountainous regions in Greece; it contributes to and increases the income of the local population, through the production of wood and other forest products. In recent decades, the cultivation of Christmas trees (CT) is an activity that has supported the development of mountainous and semi-mountainous areas of Greece.

The use of Christmas trees over the Christmas period is a Christian custom that is very broadly disseminated in Greece, and in other countries. In Greece, until 1964, the market demand for Christmas trees was covered by the felling of small trees during the clearing or thinning of forests, always in accordance with forest management plans. The increasing demand for Christmas trees and the inability to cater for this demand using domestic products resulted in the import of Christmas trees from other countries, such as Austria, Denmark and Germany (Karameris, 1996).

In response to this situation, the Forest Services took action, urging and advising farmers in mountainous and semi-mountainous regions of Greece to create artificial

In: M. Salampasis, A. Matopoulos (eds.): Proceedings of the International Conference on Information and Communication Technologies
for Sustainable Agri-production and Environment (HAICTA 2011), Skiathos, 8-11 September, 2011.

Christmas tree plantations on their land, and even provided them with subsidies and small trees for free from the state forest nurseries (Karameris, 1996).

Thus, the cultivation of Christmas trees began to develop gradually in several mountainous and semi-mountainous areas of continental Greece. The main centres involved in the production of Christmas trees are the areas of Arnaia-Polygyros, Spercheiada, Sparti, Karpenisi and Astros Kynourias (Christodoulou et al, 1991; Papaspyropoulos et al., 2008).

The production of Christmas trees in Greece mainly takes place on privately-owned land; only a very small number of trees are supplied from public forests. The main species produced include: Abies borisii Regis, Abies cephalonica, Pinus nigra, Pinus silvestris and Pseudotsouga menziesii.

The Christmas trees that come from public forests are the result of cultivation interventions made within the implementation framework of various management plans; they are also obtained from the opening of roads, from clearing lanes for electricity and telephone lines, etc.

At privately-owned areas, the production of Christmas trees takes place: a) on mountainous and semi-mountainous agricultural land that has either been abandoned or whose owners believe that the cultivation of Christmas trees is more profitable than other crops, b) at privately-owned chestnut orchards, where fir trees have been planted and c) at privately-owned forested fields, whose cultivation either for wood or fruit production necessitates the removal of certain small fir trees growing there (Karameris, 1996).

The cultivation of Christmas trees is of very high economic importance, since it creates an additional income for the inhabitants of some mountainous and semi-mountainous regions of Greece, and upgrades the natural environment (Kaloudis et al, 2002). These are areas that would have otherwise remained uncultivated and exposed to erosion; by having Christmas trees planted, these areas are covered by vegetation and flooding is prevented.

In addition, the cultivation of Christmas trees is considered to be a competitive option, compared to other agricultural crops, such as wheat, and the end product is also non-polluting (Christodoulou et al, 1991).

Nevertheless, the producers of Christmas trees face significant problems as regards the sale of their products, due to the low purchasing power of households, and because of the competition both from low-quality trees and from substitute trees (e.g. artificial Christmas trees) (Kaloudis et al, 2002).

The Greek state has not taken the necessary steps to support the cultivation of Christmas trees through subsidies; therefore, imported Christmas trees are gradually gaining points in the domestic market.

According to data from the Ministry of Environment, Energy and Climate Change, in recent decades, it has been observed that the number of felled CT is lower compared to the number of approved CT. The ratio between approved and felled CT fluctuates from year to year and also per category of origin (artificial plantations, chestnut orchards, forested agricultural lands) (MEECG, 2010).

In Europe, Denmark is one of the countries with the greatest production of Christmas trees. Its annual production is equal to 10 million trees, making it the second European country in production figures after Germany (Østergaard and Christensen, 2007). Also, in Belgium, the area covered by Christmas trees amounts to 5,000 ha and is mainly located in Wallonia (Guiot and Raymackers, 2007). In Austria, the largest share of the demand is covered by the domestic production (85%) and the rest (15%) by imports from countries such as Denmark and Germany (Schuster, 2007).

The aim of this paper is to forecast future Christmas tree production, based on the last 28 years of production (1981-2008). The two time series which were used was the number of approved trees for cutting (input value) and the number of actually cut trees (output value). The results from the model were combined with the results from an Autoregressive Integrated Moving Average (ARIMA) model and a forecast of Christmas tree usage until the year 2015 was made.

2. ARIMA Model

Univariate –ARIMA models are constructed using only the information contained in the series. Thus models are constructed as linear functions of past values of the series and/or previous random shocks (or errors). Forecasts are generated under the assumption that the past history can be translated into predictions for the future. Box and Jenkins (1976) formalized the ARIMA modelling framework by defining three steps: identification of the model, estimation of the coefficients and verification of the model. These procedures apply to stationary series (time series with no systematic change in mean and variance) whose data are normally distributed. First or second - order differences usually remedied non–stationary means, and logarithmic transformation remedied non–stationary variances and non – normal distributions of original data. Identification of the number of terms to be included in the model was based on the examination of the autocorrelation (ACF) and partial autocorrelation (PACF) functions of the difference, log–transformed time series. Estimation of the model coefficients was achieved by means of the maximum likelihood method. Verification of the model was performed through diagnostic checks of residuals (histogram and normal probability plots of residuals and standardized residuals). The simple non – seasonal ARIMA model has a general form of (p,d,q) where p is the order of the non –seasonal autoregressive term (AR), q is the order of the non – seasonal moving average term (MA) and d is the order of the non – seasonal differencing. In our case we used the approved for cutting time series an we made an estimation of approved trees for the years 2009 until 2015. The selected ARIMA model was an (1,0,0) because it presented the smaller AIC and SBC values.

3. Neural Networks

3.1. Multilayer Perceptrons

Multilayer perceptrons have been applied successfully to solve some difficult and diverse problems by training them in a supervised manner with a highly popular algorithm known as the error back-propagation algorithm. This algorithm is based on the error- correction learning rule. As such, it may be viewed as a generalization of an equally popular adaptive filtering algorithm. Basically, the error back-propagation process consists of two passes through the different layers of the network: a forward pass and a backward pass. In the forward pass, an activity pattern (input vector) is applied to the sensory nodes of the network, and its effect propagates through the network, layer by layer. Finally, a set of outputs is produced as the actual response of the network. During the forward pass the synaptic weights of the network are all fixed. During the backward pass, on the other hand, the synaptic weights are all adjusted in accordance with the error-correction rule. Specifically, the actual response of the network is subtracted from a desired (target) response to produce an error signal. This error signal is then propagated backward through the network, against the direction of synaptic connections tic weights are adjusted so as to make the actual response of the network move closer to the desired response. The error back-propagation algorithm is also referred to in the literature as the back-propagation algorithm, or simply buck-prop. Henceforth, we will refer to it as the back-propagation algorithm. The learning process performed with the algorithm is called back-propagation learning.

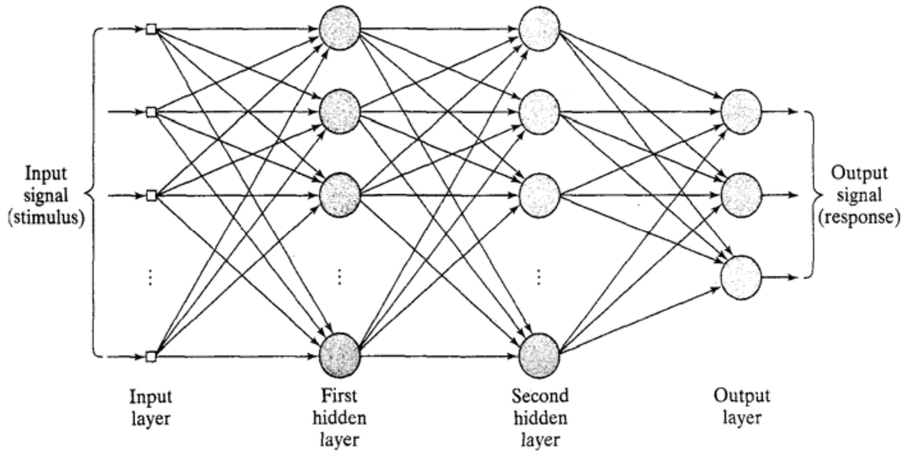

Fig. 1. A multilayer perceptron with two hidden layers

A multilayer perceptron has three distinctive characteristics:

1. The model of each neuron in the network includes a nonlinearity at the output end. The important point to emphasize here is that the nonlinearity is smooth (i.e., differentiable everywhere), as opposed to the hard-limiting used in Rosenblatt's perceptron. A commonly used form of nonlinearity that satisfies this requirement is a sigmoid al nonlinearity defined by the logistic function:

$$y_j = \frac{1}{1+\exp(-u_j)} \quad (1)$$

Where u_j is the net internal activity of the neuron and j is the output of the neuron . The presence of nonlinearities is important because, otherwise, the input-output relation of the network could be reduced to that of a single-layer perceptron. Moreover, the use of the logistic function is biologically motivated, since it attempts to account for the refractory phase of real neurons.

2. The network contains one or more layers of hidden neurons that are not part of the input or output of the network. These hidden neurons enable the network to learn com plex tasks by extracting progressively more meaningful features from the input patterns (vectors).

3. The network exhibits a high degree of connectivity, determined by the synapses of the network. A change in the connectivity of the network requires a change in the population of synaptic connections or their weights (Haykin, 1999)

3.2. Back Propagation Algorithm

The error signal at the output of neuron *j* at iteration *n* (presentation of the *n*th training example) is defined by

$e_j(n)=d_j(n) - y_j(n)$, neuron *j* is an output node

Where:

- $e_j(n)$ refers to the error signal at the output of neuron *j* for iteration *n*.
- $d_j(n)$ refers to the desired response for neuron *j*.
- $y_j(n)$ refers to the function signal appearing at the output of neuron *j* at iteration *n*.

We define the instantaneous value of the error energy for neuron *j* as $\frac{1}{2} e_j^2(n)$. Correspondingly, the instantaneous value $\varepsilon(n)$ of the total error energy is obtained by summing $\frac{1}{2} e_j^2(n)$ over all neurons in the output layer; these are the only "visible" neurons for which error signals can be calculated directly. We may thus write

$$\varepsilon(n)=\frac{1}{2}\sum_{j\in C} e_j^2(n) \quad (2)$$

where the symbol $\varepsilon(n)$ refers to the instantaneous sum of error squares or error energy at iteration *n*, the set *C* includes all the neurons in the output layer of the network. Let *N* denote the total number of patterns (examples) contained in the training set. The *average squared error energy* is obtained by summing $\varepsilon(n)$ over all *n* and then normalizing with respect to the set size *N,* as shown by

$$\varepsilon_{av} = \frac{1}{N}\sum_{n=1}^{N}\varepsilon(n) \tag{3}$$

The instantaneous error energy $\varepsilon(n)$, and therefore the average error energy ε_{av}, is a function of all the free parameters (i.e., synaptic weights and bias levels) of-the network. For a given training set, ε_{av} represents the *cost function* as a measure of learning performance. The objective of the learning process is to adjust the free parameters of the network to minimize ε_{av}. To do this minimization, we use the Levenberg–Marquardt algorithm (Haykin, 1999)

The Levenberg–Marquardt algorithm (LMA) provides a numerical solution to the problem of minimizing a function, generally nonlinear, over a space of parameters of the function. These minimization problems arise especially in least squares curve fitting and nonlinear programming.

The LMA interpolates between the Gauss–Newton algorithm (GNA) and the method of gradient descent. The LMA is more robust than the GNA, which means that in many cases it finds a solution even if it starts very far off the final minimum. On the other hand, for well-behaved functions and reasonable starting parameters, the LMA tends to be a bit slower than the GNA.

The LMA is a very popular curve-fitting algorithm used in many software applications for solving generic curve-fitting problems.

The primary application of the Levenberg–Marquardt algorithm is in the least squares curve fitting problem: given a set of empirical datum pairs of independent and dependent variables, (x_i, y_i), optimize the parameters β of the model curve f(x,β) so that the sum of the squares of the deviations

$$S(\beta) = \sum_{i=1}^{m}[y_i - f(x_i,\beta)]^2 \tag{4}$$

becomes minimal.

4. Methodology

For this paper we used the NNTool of the MatLab 2009 suite for the creation of two neural networks, the first was a Feed Forward-Back Propagation (FFBP) and the second was a Cascade – Forward Back Propagation (CFBP). In detail, we used for training a sample of 28 measurements of inputs and outputs; we used the TRAINLM function, the LEARNGD adaption learning function and the MSE, performance function. Both neural networks were composed, by 2 layers, one hidden and one output. The hidden layer transfer function was the sigmoid and the output layer transfer function was the linear. The number of neurons in the hidden layer was 70 and in the output layer was 1.

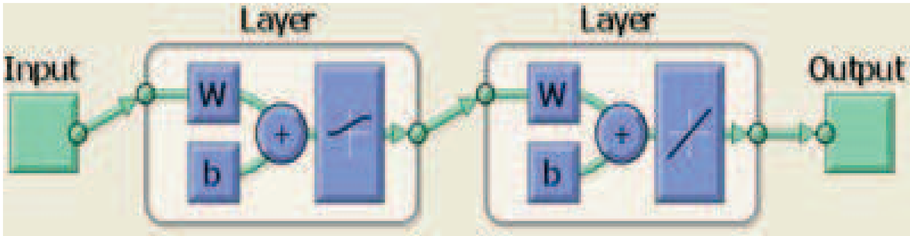

Fig. 2. The Feed forward back propagation Neural Network used for the calculations

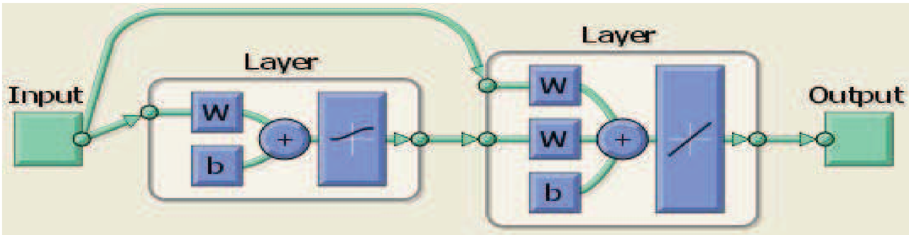

Fig. 3. Cascade-forward back propagation

In order to stop the training from over fitting we used the Validation and Testing ability of the NNTool which was supplied with inputs and outputs for better training. The networks were trained for 500 epochs.

After the initial training of the networks, we simulated the output results by providing the networks only with the inputs. The network that made the best fit to the real values (in this case the FDBP) was selected for forecasting the number of cut Christmas trees (Table 1 Results).

In doing so we provided the network with a series of measurements which were created by an ARIMA model. These measurements forecast the number of approved trees for cutting for the years 2009 to 2015.

Thus we created a hybrid model which combines the results from an ANN and an ARIMA model. Hybrid models were proved to provide better results in comparison with ANN or ARIMA models. (Koutroumanidis et al, 2009)

5. Results

The following table presents the results from the training of both ANN and the results provided by the ARIMA model. Additionally there are some statistical results which provide relative measures of accuracy.

Table 1. Model Results

	Samples	R^2
Training FDBP	21	1
Validation FDBP	4	0.87887
Testing FDBP	3	0.8775
Training CFBP	21	1
Validation CFBP	4	0.70654
Testing CFBP	3	0.41744
Feed forward back propagation (FDBP)		
sMAPE	0,050443 or 5,04 %	
NRMSD	6,40275E-06	
Cascade forward back propagation (CFBP)		
sMAPE	0,12044927 or 12%	
NRMSD	2,51239E-06	
ARIMA		
sMAPE	0,00675 or 0.67 %	
NRMSD	3,1468E-05	

R^2 is the coefficient of determination, and represents the square of the sample correlation coefficient between the outcomes and the values being used for forecasting. Values, which are closer to 1, indicate better forecasts and high degree of correlation, while values closer to 0 indicate poor forecasts and low level of correlation (Steel and Torrie, 1960; Nagelkerke, 1991).

sMAPE or symmetrical Mean Absolute Percentage Error, is an accuracy measure based on percentage (or relative) errors. It is usually defined as follows:

$$SMAPE = \frac{1}{n}\sum_{t=1}^{n}\frac{|A_t - F_t|}{(A_t + F_t)/2} \quad (5)$$

When having a perfect fit, sMAPE is zero. But in regard to its upper level the sMAPE has no restriction. The results show that for the ANN the average error of forecast is 5.04%, which represents the deviation from the real value, which will be observed.

NRMSD is the Root Mean Square Deviation. It is a frequently used measure of the differences between values forecasted by a model or an estimator and the values actually observed from the thing being modelled or estimated.

It is usually defined as follows:

$$RMSD(\theta_1,\theta_2)=\sqrt{MSE(\theta_1,\theta_2)}=\sqrt{E((\theta_1,\theta_2)^2)}=\sqrt{\frac{\sum_{i=1}^{n}(x_{1,i}-x_{2,i})^2}{n}} \quad (6)$$

$$NRMSD=\frac{RMSD}{x_{\max}-x_{\min}} \quad (7)$$

The results from the trained network and the forecast for the years 2009-2015 are presented on the following diagram.

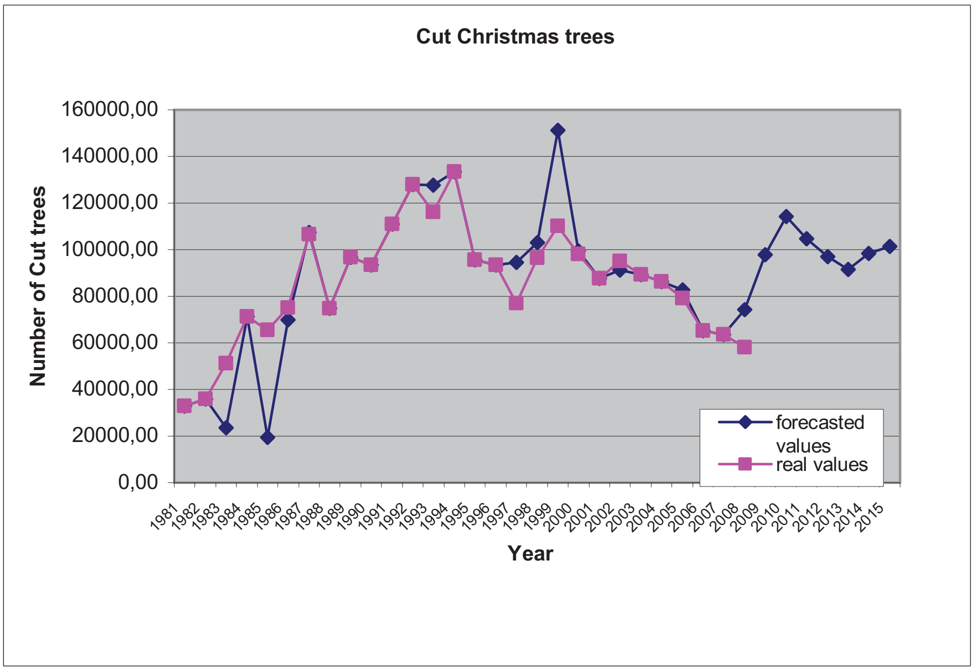

Diagram. 1. Hybrid Model Forecasts

The forecasted values were calculated from the hybrid model and are presented in diagram 1. It is obvious from the results that we expect a significant increase in the number of cut Christmas trees in the following year, followed by a decrease in the years 2010 and 2012. After the peak point which we will reach in 2010, we should expect a decrease in the number for the years 2011 to 2013.

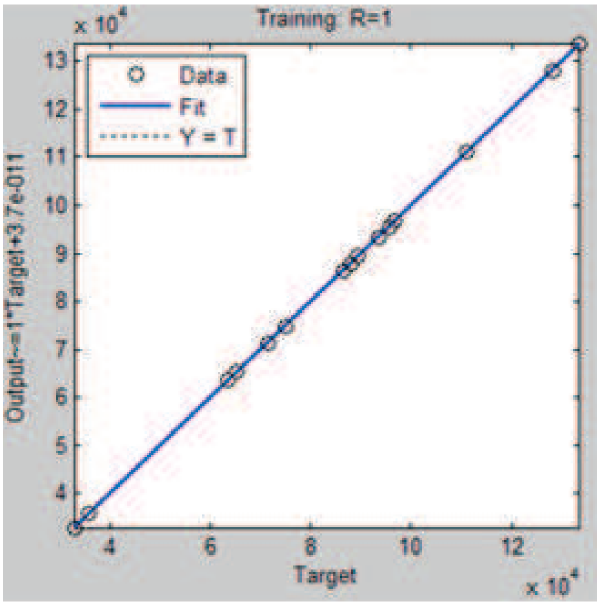

Fig. 2. Regression analysis of ANN training (FFBP)

It is obvious from diagram 2, that training results are very good; the coefficient of determination is 1.

6. Conclusions

From the above-mentioned research, it is observed that maintaining the custom of Christmas trees is a particularly profitable activity, since the production of Christmas trees offers an additional income and contributes to the development of certain mountainous and semi-mountainous rural regions, which would otherwise be destined to decline. The economic advantages involved, in combination with the environmental benefits of cultivating Christmas trees render their continued cultivation an important factor for the economy of mountainous areas and for the natural environment. In recent years, a gradual reduction of felled trees has been noted, probably due to the economic crisis, cheaper imported trees and the artificial trees on sale. Based on the results that emerge from this model, a further reduction in the number of trees that will be felled from 2011 onwards is also to be expected. This methodology provides the ability to make predictions regarding the future Christmas Trees market. These predictions could help producers to modify their production according to the future demands. In the future the predictions could be enhanced by adding more predictors (ie more time series) and / or enhancing the amount of data provided by the already selected time series. ANN's are greatly affected by the amount of input data and thus by the time series we provide them for learning and training. There are no specific disadvantages of this methodology apart from the fact that there is a need for large time series in order to provide adequate predictions.

References

1. Christodoulou, A., Blioumis, B., Papadopoulos, I. and Stamou, N. (1991) The production of Christmas trees: An attractively productive but misunderstood activity. Scientific Annals of the Department of Forestry and Natural Environment. Aristotelian University of Thessaloniki, Vol LD/1. No 11, Thessaloniki (In Greek).
2. Guiot, C. and Raymackers, D. (2008) The production of Christmas trees in Belgium. In: Thomsen, I.M., Rasmussen, H.N. & Sørensen, J.M. (Eds.): Proceedings of the 8th International Christmas Tree Research and Extension Conference. Forest & Landscape Working Papers No 26/2008: 75-76. Forest & Landscape Denmark, Hørsholm. http://www.sl.life.ku.dk/upload/working papersno26.pdf.
3. Haykin, S. (1999) Neural Networks a comprehensive foundation, Prentice Hall.
4. Kaloudis, S., Costopoulou, C., Gouskou, K. and Lorentzos, N. (2002) Electronic trading of natural Christmas trees. Geotechnical Scientific Issues 13(II/3): 22-28 (In Greek).
5. Karameris, A. (1996) The custom and cultivation of Christmas tree. Scientific Annals of the Department of Forestry and Natural Environment. Aristotelian University of Thessaloniki. Vol 39/2: 847-864. Dedicated to Dr. N. Th. Papamichos, Emeritus Professor. Thessaloniki (in Greek).
6. Koutroumanidis, T., Ioannou, K. and Arabatzis, G. (2009). Predicting fuelwood prices in Greece with the use of ARIMA models, artificial neural networks and a hybrid ARIMA–ANN model. Energy Policy, 37(9), p. 3627-3634.
7. Ministry of Environment, Energy and Climate Change, (2010) Christmas trees (artificial plantations, chestnut coppice orchards and other sources). Athens. (in Greek).
8. Nagelkerke, N. (1991) A Note on a General Definition of the Coefficient of Determination," Biometrika, 78(3), p. 691–692.
9. Østergaard, K. and Christensen, C.J. (2008). Christmas trees in Denmark. In: Thomsen, I.M., Rasmussen, H.N. & Sørensen, J.M. (Eds.): Proceedings of the 8th International Christmas Tree Research and Extension Conference. Forest & Landscape Working Papers No 26/2008: 73-75. Forest & Landscape Denmark, Hørsholm. http://www.sl.life.ku.dk/upload/working papersno26.pdf.
10. Papaspyropoulos, K., Christodoulou, A., and Blioumis, V. (2008). The status of Christmas tree market in Greece. In: Thomsen, I.M., Rasmussen, H.N. & Sørensen, J.M. (Eds.) 2008: Proceedings of the 8th International Christmas Tree Research and Extension Conference. Forest & Landscape Working Papers No

26/2008: 107-111. Forest & Landscape Denmark, Hørsholm. http://www.sl.life.ku.dk/upload/working papersno26.pdf.

11. Schuster, K. (2008). Christmas Tree Production in Austria. In: Thomsen, I.M., Rasmussen, H.N. and Sørensen, J.M. (Eds.): Proceedings of the 8th International Christmas Tree Research and Extension Conference. Forest & Landscape Working Papers No 26/2008: 78-85. Forest & Landscape Denmark, Hørsholm. http://www.sl.life.ku.dk/upload/working papersno26.pdf.

12. Steel, R. G. D. and Torrie, J. H., (1960) Principles and Procedures of Statistics, New York: McGraw-Hill, pp. 187, 287.