http://www.pjbs.org

PJBS

ISSN 1028-8880

Pakistan Journal of Biological Sciences

ANSInet
Asian Network for Scientific Information
308 Lasani Town, Sargodha Road, Faisalabad - Pakistan

Pakistan Journal of Biological Sciences 9 (15): 2841-2846, 2006
ISSN 1028-8880

Classification of Hyperspectral Data and Neural Networks to Differentiate Between Typical Leaves of Wheat and Those Deficient in Nitrogen, Phosphorus, Potassium and Calcium

[1]Tomas Ayala-Silva, [2]Caula A. Beyl and [1]Robert R. Heath
[1]USDA-ARS National Germplasm Repository, Subtropical Horticulture Research Station, 13601 Old Cutler Road, Miami, FL 33158, USA
[2]Department of Plant and Soil Science, Alabama A and M University P.O. Box 1208, Normal, AL 35762, USA

Abstract: A fast identification of insufficiency of nutrients using spectral features would be a useful instrument in farming and in other nutrient demanding agricultural systems such as those proposed for long period space missions. A Multilayer Perceptron (MLP) neural network and backpropagation algorithm was used to differentiate between normal leaves of wheat (*Triticum aestivum* L.) and those deficient in nitrogen, phosphorus, (K) and (Ca) using hyperspectral data. The network consisted of three layers with spectral reflectance of the leaves in wavelengths from 401 to 770 nm as the input layer and the nutrient concentrations as the output layer. Based upon the values of actual nutrient concentrations (mg L^{-1}), plants were classified as either deficient or standard. Wheat plants were grown for ≈100 days under both hydroponic conditions in the greenhouse and vermiculite media in a growth chamber using Hoagland's complete nutrient solution with selected minerals eliminated to induce specific nutrient deficiencies. Check plants received complete nutrient solutions. The MLP model was trained and tested successfully within 1000 epochs as the MSE of the sample-training curve approached zero. The backpropagation algorithm functioned well with the following accuracies for the classification model: N 90.9, P 100, K 90 and Ca 100%. Using the regression model, the following accuracies were obtained: N 93.0, P 87.2, K 91.9 and Ca 97.4%. This affirms the potential of using spectral data coupled with either a classification or regression neural network models to quickly categorize leaves deficient in these four major minerals so that remedial applications of those nutrients can be made before the yield is drastically affected.

Key words: Deficiency, neural network, reflectance, hyperspectral data, nutrients

INTRODUCTION

The Human Exploration and Development of Space (HEDS) Division of NASA has developed systems for the production of edible crops to be used as food during long-term space missions and colonization missions. An ideal controlled environment life-support system (CELSS) is bioregenerative and possible during space missions when resupply from Earth is very costly or unfeasible. Through biological processes (photosynthesis, transpiration), plants in a CELSS environment absorb CO_2, release O_2, produce edible food and recycle water (Bugbee and Salisbury, 1987). For a CELSS structure, with its emphasis on space, labor, equipment and weight reduction, real time comprehensive crop-response models are needed (Frick *et al*., 1998). These models will be needed to predict precise responses to changes in the environment and crop needs (i.e., nutrient deficiencies or water deficit). Wheat, the focal point of this study, is one of the crops considered in the CELSS program.

One major factor that could hinder the growth of plants in a structure such as this is the accessibility to nutrients. If leaf spectroreflectance can be used to detect specific nutrient stresses, it would enable timely detection of deficiencies and allow precision supplementation to prevent loss of production without the costs innate in replacement of spent nutrient solutions. Space plant production systems will not be able to support large amounts of solid materials because of the costs of launching large quantities of mass (including water) into orbit (Wheeler *et al*., 1990).

Estimated or reliable crop responses are frequently difficult to obtain. In this study, a simple Artificial Neural Network (ANN) was developed as a substitute to the

Corresponding Author: Tomas Ayala-Silva, USDA-ARS National Germplasm Repository, 13601 Old Cutler Road, Miami, FL 33158, USA Tel: (305) 254-4484 Fax: 305-969-6410

more common approach using a statistical regression model to describe the spectral response of wheat to various elemental deficiencies and their effects on spectral reflectance.

Neural network systems are changeable systems that can learn associations through repetitive presentation of data. Artificial neural networks (ANN) have a wide variety of applications (Principe *et al*., 2000) including classification of various types of data to accuracies that are generally equivalent or higher than those derived from conventional statistics (Hepner *et al*., 1990; Medina and Vasquez, 1991). In essence, an ANN may be considered a large number of simple unified neurons or units that work in parallel to sort input data into output classes (Aleksander and Morton, 1995; Hepner *et al*., 1990). The interconnections between the units are weighted and, with the input data, these weights determine the level of creation of a unit in the network, which in turn influences the echelon of activation of other units in the network and eventually determines the network outputs. The magnitude of the weights is determined by an interactive training procedure through which the network repeatedly tries to learn the correct output for each of the training samples. The procedures involve modifying the weights between units until the artificial neural network is able to differentiate the training data correctly.

A backpropagation-learning algorithm has been commonly used in which the error between the network output and desired output is minimized (Haykin, 1999). Once trained, the ANN may then be used to determine class association for other data. The network design parameters (i.e., data set size, weights, training and testing sets) influence the performance of the ANN but may be difficult to define and are characteristically determined instinctively (Ripley, 1996). Once trained and defined, however, the ANN is more dynamic than standard statistical analyzers. Artificial neural networks are more tolerant to noise and missing data, can adapt over time, evaluate the importance of data in the analysis and once trained, can be more efficient computationally than other orthodox analyzers (Swingler, 1996). As with conventional numerical classifiers, the characteristics of the training set, for instance, may be very important in determining the relative precision of the analysis derived from ANN and orthodox statistics. The system we describe, Neuro Dimension, Inc. (1998) has the capabilities of generalizing to new, unseen data. Neural networks can be used for both regression and classification. The objective of this study was to develop and evaluate both regression and classification neural network models to use hyperspectral data for detection and bias between individual nutrient deficiencies.

MATERIALS AND METHODS

Greenhouse: Five wheat seeds were planted into each 1.5×1.5×1.5 cm horticube (Smithers-Oasis, Kent, OH, USA). The plants were grown in the greenhouse under a 14/10-day/night cycle with a 25/21°C day/night temperature and relative humidity of 70/55% day/night during fall, spring and winter and a 30/25°C day/night temperature during the summer. When the seedlings had reached 5 to 6 cm in height, the seedlings were transferred three per/hole into foam plugs in the hydroponic system, consisting of two polyvinyl carbonate (PVC) pipes, 10 cm in diameter and 100 cm long, connected at each end to 2 cm PVC pipes. The hydroponic system had five holes, 5 cm in diameter and spaced 17 cm apart on top of a 105 cm long, 10 cm diameter PVC pipe. To keep the liquid medium circulating and the root system moist, at one end of the system was attached a submersed pump (Little Giant 2E-38N, Oklahoma City, OK, USA) controlled by a timer. The pumps ran continuously for one h, every four h with a flow rate of 1 L min^{-1}. When the pumps were off, the solution flowed back down the fill/drain fittings. The amount of solution that remained in the system was controlled by overflow fittings (located at the input and output entrance of each system) 2 cm in diameter allowing a depth of six cm of solution. All systems were connected to air pumps providing continuous airflow through the nutrient medium.

Growth chamber: Seeds of wheat *Triticum aestivum* L. 'Pioneer 2693' (Pioneer Hi-Bred International Inc., Des Moines, IA, USA) were rolled in paper towels, kept moist with deonized water for 24 h in Petri dishes and then transferred to 10×10×12 cm green plastic containers (Hummert International, Earth City, MO, USA) containing fine grade vermiculite (Grace, Milpitas, CA, USA) that was fully hydrated with water. Six seeds per container were placed about 3 cm beneath the surface of the vermiculite. Seedlings were thinned to three plants/pot after two weeks. The plants were grown in an environmentally controlled 3.0×2.0×2.5 m walk-in growth chamber (Rheem Puffer Hubbard Environmental TM, NY, NY, USA), under a mixture of fluorescent and incandescent lighting on a 14/10 day/night cycle with a photosynthetically active radiance of 250-340 μE m^{-2} S^{-1}, 25/21°C day/night temperature and relative humidity of 70/55% day/night during the spring, fall and winter and a 30/25°C day/night temperature during the summer.

Induction of deficiency and analysis: Selected nutrient deficiencies were forced after the initial 20 days of growth. The growth solution consisted of a 100% Hoagland's

Table 1: Hoagland nutrient solution with selected components eliminated (- omitted, + included) for induction of deficiency in nitrogen (N), phosphorus (P), potassium (K) and calcium (Ca)

Components of growth solution	Desired mineral (salt) deficiency treatment				
	Complete	-N	-P	-K	-Ca
NH_4NO_3	+	-	+	+	+
KH_2PO_4	+	+	-	-	+
$CaCl_2$	+	+	+	+	-
$MgSO_4 \cdot 7H_2O$	+	+	+	+	+
KNO_3	+	-	+	-	+
Fe(II)-EDTA	+	+	+	+	+
H_3BO_4	+	+	+	+	+
$ZnSO_4 \cdot 7H_2O$	+	+	+	+	+
$CuSO_4 \cdot 5H_2O$	+	+	+	+	+
$MnSO_4 \cdot H_2O$	+	+	+	+	+
$Na_2MoO_4 \cdot 2H_2O$	+	+	+	+	+
$CoSO_4 \cdot 7H_2O$	+	+	+	+	+

nutrient solution (Hoagland and Arnon, 1950) with selected constituents eliminated in certain treatments to induce specific elemental deficiencies (Table 1). The check basal nutrient solution contained the following: 50 μM NH_4, 10 μM P, 2200 μM K, 1000 μM Ca, 500 μM Mg, 500 μM S, 50 μM Cl, 12.5 μM Bo, 0.1 μM Mo. The micronutrient metals Fe, Zn, Cu, Mn, were added as EDTA (ethylenediaminetetraacetate) chelates. The check solution contained at least: 20 μM Fe, 5 μM Zn, 3 μM Cu and 0.3 μM Mn. The solution was maintained at pH 6.0±0.2 by measuring every three days and adjusting with 1 N H_2SO_4 or 1 N NaOH.

When the plants first began exhibiting visual symptoms of nutrient deficiency at about ≈9 weeks, data were collected on height, spectral reflectance (250 to 1100 nm), chlorophyll content and visual appearance.

Chlorophyll concentration was determined using the Minolta Chlorophyll Meter (SPAD-502; Spectrum Technologies Inc., Plainfield, IL, USA). The chlorophyll (in SPAD units) was converted to actual chlorophyll concentrations in mg g^{-1} of fresh leaf tissue using a derived regression equation (Ayala-Silva and AL-Hamdani, 1997). To develop the equation, chlorophyll was extracted using dimethylformamide (DMF) and analyzed using the method of Ayala-Silva and AL-Hamdani (1997) and then absorbance measured at 647 and 664 nm using a Spectronic 601 spectrophotometer (Milton Roy, Rochester, NY, USA). Spectral reflectance between 250 and 1100 nm were taken from the attached leaves prior to harvest, using a spectroradiometer (GER 1500 System, Geophysical and Environmental Research Corporation, Milbrook, NY, USA) equipped with an optical cable. Reflectance of wheat leaves was obtained in the interveinal area near the midrib. In the greenhouse, data were collected between 1100 and 1300 h to minimize the impact of changing sun angle.

Reflectance was calculated as ratio of reflected to incident radiation, as determined by frequent measurement of reflected radiation from a calibrated white plate (Geophysical and Environmental Research Corporation, Milbrook, NY, USA). After 90 days of treatment, spectral data were collected, plants were harvested and the shoots were dried at 80°C for 48 h. The shoot tissue was ground with a Foss Tecator 1093 sample mill (Foss Tecator, Hoganas, Sweden) to pass a 20-mesh sieve and digested for determination of N, P, K and Ca (Hu and Barker, 1999).

Nitrogen analysis was performed using the LECO FP-2000 Nitrogen Analyzer Determinator (LECO Corporation, St. Joseph, MI, USA). Macronutrient contents were determined using inductively coupled plasma (ICP, Plasma 400 Analysis Version 4.10; Perkin Elmer Inc., Norwalk, CT, USA). Plants were then classified as either deficient or not deficient based upon expected normal values for each nutrient (Jones *et al*., 1991).

Data analysis: Treatments were arranged in a completely randomized block design with two blocks and ten sub samples for each of the five elemental treatments. The locations (greenhouse and growth chamber) were analyzed separately using conventional statistics because rates of growth and manifestation of symptoms differed in the two environments. Data were statistically analyzed using SAS procedures and reported at the 0.05 level of significance. Means were separated using Tukey's Studentized Range (HSD, $p<0.05$) test. Statistical models (ANOVA and means separation) were performed using the GLM procedures of PC-SAS Version 8.0 (SAS Institute, 1999). A correlation analysis (data not shown) using SAS was done to determine if the reflectance at each wavelength was positively or negatively correlated with actual tissue nutrient and chlorophyll concentrations.

A Multilayer Perceptron (MLP) neural network and backpropagation algorithm was used to distinguish between check leaves of wheat (*Triticum aestivum* L.) and those deficient in nitrogen (N), phosphorus (P), potassium (K) and calcium (Ca) using hyperspectral data. The network (Fig. 1) consisted of a simple, fully connected, multilayer perceptron model with one-hidden layer (NeuroSolutions, Inc., Gainesville, FL, USA). Even though spectral reflectance was measured from 250-1000 nm, only wavelengths from 401-770 nm were used in the neural network model. These wavelengths were selected based on the correlations obtained between the wavelengths and the deficiencies. The reflectance of these wavelengths served as the input layer and the output layer consisted of either (1) nutrient deficiency classes -N,

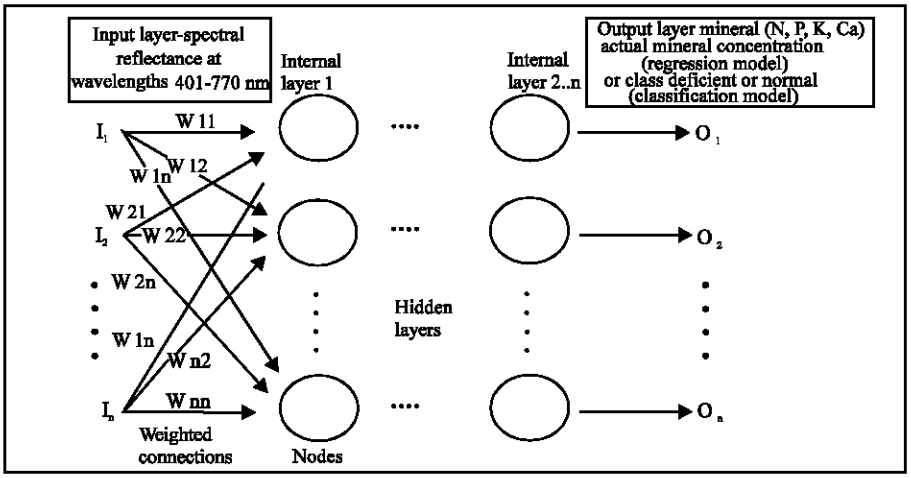


Fig. 1: The neural network structure with one hidden layer. The input layer (I_I) consists of spectral reflectance from 401 to 770 nm. The inputs are connected to hidden nodes, which are in turn connected to the output layer (O_n)

-P, -K, -Ca in the classified model or (2) the actual tissue concentration of the four macronutrients (N, P, K and Ca) in the regression model. The number of processing elements used was a function of input channels and the number of exemplars. Data were split into two sets: The training set for model building (5 sub samples) and the validation (test) set (data from greenhouse and growth chamber pooled).

RESULTS AND DISCUSSION

The MLP model was trained and tested successfully within 1000 epochs as the MSE of the sample training curve approached zero (Fig. 2). Based upon the measured values of actual nutrient concentrations (mg L^{-1}), plants were classified as either deficient or normal. Data collected from the greenhouse and the growth chamber were used to develop the neural network models. As with orthodox statistical analysis, the characteristics of the training data influence classification performance. The size of the training set, for example, may be important in determining the accuracy of classification or regression derived from AANs or orthodox statistics (Hepner *et al*., 1990). The data obtained from the greenhouse and the growth chamber were pooled for the neural network analysis. The network performed better using the larger number of measurements (pooled data) and creating one model which could then be used irrespective of environment was potentially very valuable. An ANN was developed as a substitute to the traditional statistical model. The resulting neural network classification model using the backpropagation algorithm correctly identified plants deficient in nitrogen (N) 90.9, phosphorus (P) 100, potassium (K) 90 and calcium (Ca) 100% (Table 2) based on input of reflectance data from 401 to 770 nm. With the regression model, the following accuracies were obtained: N 93.0, P 87.2, K 91.9 and Ca 97.4% respectively (Fig. 3). Beyl *et al*. (2000) demonstrated the efficiency of a similar type of neural network model to differentiate between species (tomato, lettuce and wheat) using spectral reflectance as a source.

Table 2: Network output and performance using the classification (a) and regression (b) model for wheat shoots deficient in nitrogen (N), phosphorus (P), potassium (K) and calcium (Ca)

(a)

Output/Desired	N	P	K	Ca
N	30.0000	0.0000	0.0000	0.0000
P	3.0000	34.0000	3.0000	0.0000
K	0.0000	0.0000	27.0000	0.0000
Ca	0.0000	0.0000	0.0000	29.0000

(b)

Performance	N	P	K	Ca
MSE	0.0287	0.0512	0.0361	0.0112
Min Abs Error	0.0000	0.0049	0.0033	0.0002
Max Abs Error	0.0099	0.9672	0.5842	0.4523
r	0.9302	0.8721	0.9186	0.9738
Percent Correct	90.9091	100.0000	90.0000	100.0000

Table 3: ANOVA for wheat shoots grown in nitrogen deficient media

Source	df	Mean square	F-value	Pr>F
Replicate	7	1168.19	1.21	0.3194
Treatment	6	4.455476	3.01	0. 0155*
Location	1	656.57	0.74	0.3942

* indicates significant at $p<0.05$

The analysis of variance (ANOVA) for reflectance of nitrogen (Table 3) deficient plants showed that only treatment had a significant effect on nitrogen concentration, however, for phosphorus, potassium and

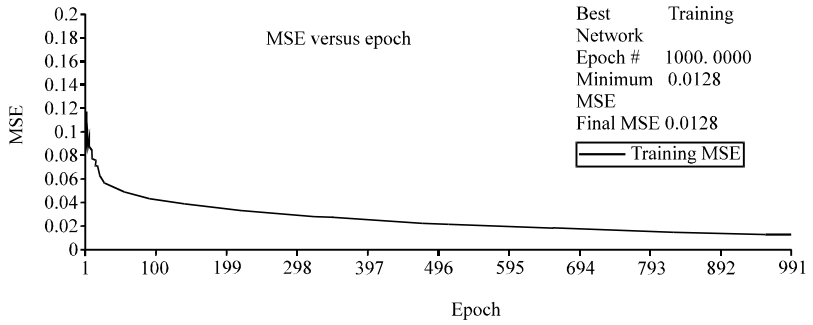
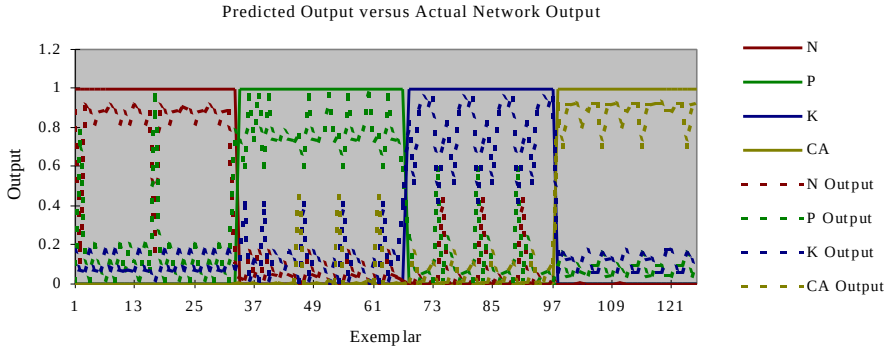


Fig. 2: Reduction in training Mean Square Error (MSE) of the best neural network model at 1000 epochs and MSE 0.0128



Performance	N	P	K	Ca
MSE	0.0287	0.0512	0.0361	0.0361
Min Abs Error	0.0099	0.0049	0.0033	0.0033
Max Abs Error	0.9302	0.9672	0.5842	0.5842
r	90.91	0.8721	0.9186	0.9186

Fig. 3: Performance of the neural network regression model for wheat shoots deficient in nitrogen, phosphorus, potassium, and calcium. Solid lines represent actual concentration and dashed lines predicted

Table 4: ANOVA for wheat shoots grown in phosphorus deficient media

Source	df	Mean square	F-value	Pr>F
Replicate	7	0.9415	0.91	0.5077
Treatment	6	531.607	514.15	0.0001**
Location	1	0.039	16.46	0.0002**

** Highly significant at (p<0.001) respectively

Table 5: ANOVA for wheat shoots grown in potassium deficiency

Source	df	Mean square	F-value	Pr>F
Replicate	7	2.26	2.98	0.0125
Treatment	6	682.95	899.91	0.0001**
Location	1	27.1	281.13	0.0001**

** Highly significant at (p<0.001) respectively

Table 6: ANOVA for wheat shoots grown in calcium deficiency

Source	df	Mean square	F-value	Pr>F
Replicate	7	2.7	0.76	0.6205
Treatment	6	274.6	77.42	0.0001**
Location	1	8.95	43.64	0.0001**

** Highly significant at (p<0.001) respectively

calcium (Table 4-6) both location and treatment were highly significant. Wavelengths from 474 to 701 nm (data not shown) were correlated to phosphorus deficiencies, whereas, potassium was highly correlated at wavelengths from 416 to 720 nm. On the other hand calcium and nitrogen were less correlated with correlations ranging from 416 to 532 and 518 to 570 nm, respectively.

The ability of the network to distinguish among the four deficiencies illustrates the utility of the application as an addition to orthodox statistics. The fact that the minerals could be effortlessly identified from the information also suggests that a common ANN could be built to relate nutrient matter with spectral signatures independent of deficiency since the model can differentiate minerals effects effortlessly.

REFERENCES

Aleksander, I. and H. Morton, 1995. An Introduction to Neural Computing. 2nd Edn., International Computer Press, Boston, MA, USA.

Ayala-Silva, T. and S. AL-Hamdani, 1997. Interactive effects of polylactic acid with different aluminum concentrations on growth, pigment concentrations and carbohydrate accumulation of Azolla, Am. Fern. J., 87: 120-126.

Beyl, C.A., R.O. Pacumbaba Jr., C. Cooper and T. Ayala-Silva, 2000. Discrimination of species and age of leaf by spectroreflectivity and neural network analysis. In Proceeding of the World Automation Conferance, Maui, Hawaii (WAC 2000), June 2000.

Bugbee, B. and F.B. Salisbury. 1987. Physiological and genetic studies on wheat for the controlled environments of the space. OSSA, Space Life Sciences Symposium, pp: 246-248.

Frick, J., C. Precetti and C.A. Mitchell, 1998. Predicting Lettuce Canopy Photosynthesis with Statistical and Neural Network Models. J. Am. Soc. Hortic. Sci., 123: 1076-1080.

Haykin, S., 1999. Neural Networks, A Comprehensive Foundation 2nd Edn., Prentice Hall, New Jersey, USA.

Hepner, G.F., T. Logan, N. Ritter and N. Bryant, 1990. Artificial neural network classification using a minimal training set: Comparison to conventional supervised classification. Photogrammetric Engineering and Remote Sensing, 56: 469-473.

Hoagland, D.R. and D.I. Arnon, 1950. The water culture method for growing plants without soil. Calif. Agric. Exper. Sta., Circ. No. 347. Berkeley, CA, USA.

Hu, Y. and A. Barker. 1999. A single plant tissue digestion for macronutrient analysis. Commun. Soil Sci. Plant Anal., 30: 677-678.

Jones, J.B., Jr, B. Wolf and H.A. Mills, 1991. Plant Analysis Handbook II: a Practical Sampling, Preparation, Analysis and Interpretation Guide. Athens (GA): Micro-Macro Publishing, pp: 130.

Medina, F.I. and R. Vazquez, 1991. Use of simulated neural networks for aerial image classification, ACSM-ASPRS Annual Convention, Bethesda, MD, pp: 264-268.

Neuro Dimension, Inc., 1998. Neuro Solutions. Version 3.0. Gainesville, Florida, USA.

Principe, J.C., N.R. Euliano and W. Curt Lefebvre, 2000. Neural and Adaptive Systems: Fundamentals Through Simulation. John Wiley and Sons, Inc., NY, USA.

Ripley, B.D., 1996. Pattern recognition and neural networks. Cambridge University Press, NY,USA.

SAS Institute, 1999. SAS/STAT User's Guide, Version 8.0. SAS Institute Inc., Cary, NC.

Swingler, K., 1996. Applying neural networks. A Practical Guide. Morgan Kaufmann Publishers, Inc., San Francisco, CA,USA.

Wheeler, R.M., C. L. Mackowiak, J.C. Sager, W.M. Knott and C.R. Hinkle, 1990. Potato growth and yield using NFT. Am. Potato J., 67: 177-187.