YYÜ TAR BİL DERG
(YYU J AGR SCI)
2015, 25(3): 285-292

Geliş Tarihi (Received): 20.04.2015
Kabul Tarihi (Accepted): 08.09.2015

Araştırma Makalesi/*Research Article (Orjinal Paper)*

The Prediction of Saint John's Wort Leaves' Chlorophyll Concentration Index using Image Processing with Artificial Neural Network

Mehmet Serhat ODABAS[1*], Sreekala BAJWA[2], Chiwan LEE[3], Erdem Emin MARAŞ[4]

[1]:Ondokuz Mayis University, Bafra Vocational School, Samsun, Turkey
[2]:North Dakota State University, Department of Agricultural and Biosystems Engineering, Fargo, ND, United States
[3]: North Dakota State University, Department of Plant Science, Fargo, ND, United States
[4]:Ondokuz Mayis University, School of Aviation, Samsun, Turkey
e-mail*: mserhat@omu.edu.tr. tel:+903121919/7534, Fax:+903625426761

Abstract: There are several methods for detecting plant nitrogen content including plant analysis like leaf chlorophyll measurement and remote sensing techniques. In this study, image-processing method was used to predict St. John's wort (*Hypericum perforatum* L.) leaf chlorophyll concentration from leaves. The experiment was carried out in greenhouse conditions. The Hougland solution was used as a fertilizer. It was applied at 5 different levels to the St. John's wort grown in pots. SPAD-502 chlorophyll meter was used for measuring the chlorophyll concentration of the leaves. The chlorophyll-a (chl-a) and chlorophyll-b (chl-b) of the leaves were measured by UV spectrometer. Artificial Neural Network (ANN) model was developed based on the RGB (red, green, and blue) components of the color image captured with a digital camera for estimating the chlorophyll concentration. According to the obtained results, the neural network model is capable of estimating the St. John's wort leaf chlorophyll concentration with a reasonable accuracy. The coefficient of determination (R^2) was 0.99 and mean square error (MSE) was obtained 0.005 from validation.

Key words: Image processing, Artificial neural network, Chlorophyll, *Hypericum,* St. John's Worth

Yapay Sinir Ağı ile Görüntü İşleme Kullanarak Kantaronda Klorofil Konsantrasyon Endeksi Tahmini

Özet: Bitki azot içeriğinin tespiti için bitkisel analizlerin dahil olduğu yaprak klorofil ölçümü ve uzaktan algılama tekniklerin de dahil olduğu çeşitli yöntemler mevcuttur. Bu çalışmada, görüntü işleme yöntemi ile kantaron (*Hypericum perforatum* L.) yapraklarının klorofil konsantrasyonu tahmin edilmiştir. Araştırmada, saksılarda yetiştirilen kantaronlara 5 farklı dozda Hougland solusyonu gübre olarak uygulanmıştır. Yaprakların klorofil konsantrasyonunun ölçülmesinde SPAD-502 klorofil metre kullanılmıştır. UV spektrometresi ile yaprakların klorofil-a (CHL-a) ve klorofil-b (CHL-b) içerikleri ölçülmüştür. Yapay Sinir Ağı (YSA) modeli kullanılarak klorofil konsantrasyonunu tahmin etmek için bir dijital kamera ile çekilen renkli görüntülerin RGB (kırmızı, yeşil ve mavi) bileşenlerinden faydalanılmıştır. Sonuç olarak yapay sinir ağı ile yüksek doğrulukta kantaron yapraklarının klorofil konsantrasyonunu tahmin edilmiştir. Doğrulama R^2 0.99 ve MSE 0.005 olarak elde edilmiştir .Bu değerler yapay sinir ağı modelinin güvenirliliğini ortaya koymaktadır.

Anahtar kelimeler: Görüntü işleme, Yapay sinir ağı, Klorofil, *Hypericum,* Kantaron

Introduction

Hypericum is a genus of about 400 species of flowering plants in the family of *Guttiferae* and the species of this genus have been preferred as traditional medicinal plants due to their wound-healing, bactericide, anti-inflammatory, diuretic and sedative properties for hundreds of years. In particular, extracts of St. John's wort (*Hypericum perforatum* L.) leaves are now widely used in Europe as a drug for the treatment of depression (Cirak et al. 2011). The leaves are very important part of plants in concerning some physiological phenomenon. light, photosynthesis, respiration, plant water consumption and transpiration.

In addition, leaves play an important role in some cultural practices such as irrigation, fertilization, etc. (Odabas et al. 2014a).

Traditionally, fertilizers are applied onto the whole field regardless of the variation across the land (Temizel et al. 2014). There are several methods for detecting plant nitrogen content including plant analysis, leaf chlorophyll measurement, and remote sensing systems. These methods are very time consuming and expensive approaches and not suitable for nitrogen applications at different rates (Noh et al. 2006). Leaf chlorophyll measurement is a useful tool for predicting crop nitrogen status (Odabas et al. 2014b).

To obtain more reliable information on nitrogen status for increasing the effectiveness of nitrogen management, researchers developed various methods to estimate leaf chlorophyll content on the basis of spectral images of the crop canopy (Namrata et al. 2007). Image processing has been used reliably in crop management and detection of nitrogen stress (Koumpouros et al. 2004; Pydipati et al. 2006). The development of low-cost digital cameras enabled the capture of images for estimating plant nitrogen status. This technology was developed for the first time by Woebbecke et al. (1995) to split images into different areas, representing soil and plants. Moreover, Liangliang et al. (2004) and Pagola et al. (2009) successfully developed a conventional digital camera for assessing chlorophyll content of winter wheat and barley, respectively.

Selection of an appropriate technique very important in obtaining a robust prediction model for leaf chlorophyll content. Artificial neural networks have been widely applied to non-linear models (Chen et al. 2007). Chen et al. (2007) developed an ANN model for predicting the chlorophyll content of cabbage seedlings. Their results indicated that the ANN model could be used to develop a practical remote sensing system with reasonable accuracy (R2 = 0.93).

The traditional chlorophyll analysis methods used in laboratory and some chlorophyll measurements also have been developed to measure chlorophyll content in leaves under field conditions (Markwell et al. 2005). Unfortunately, traditional methods require people in physical contact with plant leaves, and tend to be labor and time-consuming, limiting their application in large fields. To monitor the leaf chlorophyll content in the large fields, the using camera has been considered as a practical approach for a long time (Chappelle et al. 1992). These acquired images can provide aerial views of the field such as the chlorophyll content of the crops, can be identified, categorized, and measured using various computer technologies. For example, Kawashima and Nakatani (1998) used a portable color video camera and a personal computer to estimate the chlorophyll content in leaves.

The present research suggested an ANN model to estimate St. John's wort leaf chlorophyll, based on image processing captured with a conventional digital camera.

Materials and Methods

This experiment was carried out in North Dakota State University Greenhouse, United States in October 2013 – April 2014. The St John's wort seeds were planted in 20 plastic pots of 22 cm diameter and 25 cm height filled with soil mediums (Sunshine Mix #2 and Sunshine #3). The Sunshine Mix #2 (no base fertilizers) was used as a soil medium for seed germination. It is formulated with Canadian Sphagnum peat moss, coarse perlite and dolomitic limestone and highly recommended for cutting propagation, bedding and vegetable plants, and pot crops and seed germination. Germinated seeds were transplanted into pots filled with Sunshine #3. This mix is suitable for a wide range of growing needs, particularly where less frequent watering is desired. It has a shorter fiber peat than our other mixes, and a low fertilizer charge to protect seedlings and plugs

Standard Hoagland solution was used as fertilizer. The Hoagland solution provides all necessary nutrients for plant growth and is appropriate for growing large variety of plant species. Plants were watered everyday with the solution. In the experiment, five different doses of solution (check, quarter of the solution, half of the solution, standard solution, and two times more of the solution) were used and each treatment was replicated 4 times.

Table1. Standard Hoagland solution used the experiment.

	meq/L	mg/L	g/120 L
NH_4NO_3 (mw 80)	2	160	19.2
$Ca(NO_3)_2 \cdot 4H_2O$ (mw 236, eq 118)	5	590	70.8
KNO_3 (mw 101)	4	404	48.5
KH_2PO_4 (mw 120, eq 120)	2	240	28.8
$MgSO_4 \cdot 7H_2O$ (mw 246, eq 123)	2	246	29.5
$NH_4H_2PO_4$ (mw 115. eq 115)	2	230	27.6

The air temperature and the relative humidity of the chamber were adjusted to 23 °C, 20 °C and 40%, 50% for day and night, respectively. Konika Minolta SPAD-502 chlorophyll meter was used to measure the leaf chlorophyll concentration index. Varian Cary 50Bio UV-Spectrometer (190 -1100 nm wavelength) was used to measure the chlorophyll-a (Chl-a) and chlorophyll-b (Chl-b). Images were taken from each sampled leaf by a digital camera (DSC-TX7 Sony Corp., Tokyo, Japan).

Leaves were illuminated with incandescent light and photographed from 30 cm distance at a lens setting of 9.7 mm focal length. The resolution of the images was 2048 × 1536 pixels and they were recorded in jpeg format. The images consist of 3 components, red (R), green (G), and blue (B), and each component had 256 graduations. The image processing was performed by MATLAB software (Matlab ® R2014a (8.3.0.532) 64-bit). 1500 images were acquired in this experiment.

Chlorphyll analysis

Leaves (5 g fresh weight for each sample) were crushed by adding 80% Acetone. Each sample was homogenized in 5 cm^3 of cold acetone through precise manual creaming of the plants in cooled porcelain mortar. Obtained suspension was absorption spectra of chlorophyll extract were taken with a UV spectrophotometer and the 80 % acetone solution as the reference, at the temperature of 21 ^{0}C. Absorbance of the supernatant was measured at 663.8 nm and 646.8 nm with UV spectrophotometer. Chl-a and Chl-b concentrations (μg ml^{-1}) were calculated using extinction coefficients and simultaneous equations (eq. 1 and eq. 2) by Wellburn (1994).

$[Chl\text{-}a] = 12.19 \times E_{665} - 3.45 \times E_{649}$ ……………………….. (eq. 1)

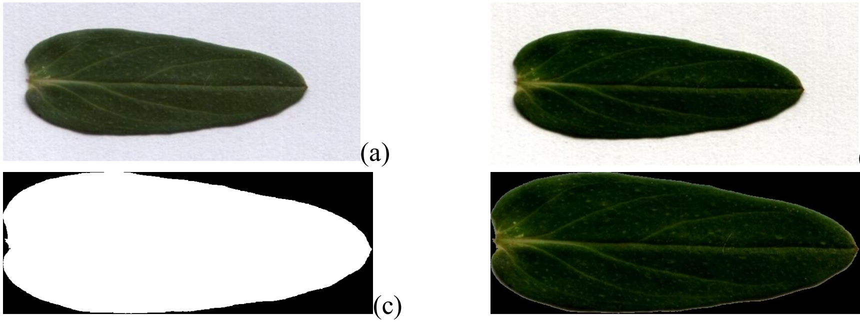
$[Chl\text{-}b] = 21.99 \times E_{649} - 5.32 \times E_{665}$ ……………………….. (eq. 2)

Image processing

Image processing is the application of computer processing techniques on the images made available to it through relevant input mechanisms. Image processing includes acquiring input associated with a digital image, changing the image to suit our needs, acquiring information from the image, which is relevant to us, and producing a desired output.

Image processing

The image processing was carried out by separating the original image of individual leaf into 3 monochrome images including red (R), green (G), and blue (B). The reflectance of the St. John's wort leaf was stronger than its background in images (Figure 1).



(a) (b) (c) (d)

Figure 1. Step of image processing a) Original leaf image, b) Normalized image, c) Crop/masked image, and d) Noise cancelled image

Therefore, the following threshold function was employed to segment the leaf from the background

$$g_i(x,y)=\begin{cases}0 & f_i(x,y)\le T_i\\ f_i(x,y) & f_i(x,y)\ge T_i\end{cases} \quad \text{……………………..(eq. 5)}$$

Where g_i (x, y) is the segmented gray level at pixel (x, y), f_i (x, y) is the original gray level at pixel (x, y), T_i is a threshold value and *i* represent the RGB components. This equation 5 (eq. 5), the pixel values of the leaf are not changed and the pixel values of the background equal zero. In the next step, the St. John's wort leaf reflectance was calculated on the basis of average intensity values of the monochrome images. The average intensity values were calculated as:

$$A_i=\frac{\sum g_i(x,y)}{N_i} \quad \text{……………………………… (eq. 6)}$$

Where A_i is the average intensity value at component i, g_i (x, y) is the intensity value of the pixel representing St. john's wort leaves, N_i is the total number of pixels representing St. John's wort leaves in component *i*, and *i* represents the RGB components of the image.

In most testing samples, the prediction errors of chl-b were higher than chl-a. The reason is that the global threshold values used at the step of the coarse extraction of green pixel are inclined to blue-green rather than yellow-green. Due to chl-a and chl-b being blue-green and yellow-green, respectively (Palta, 1990), the prediction performances for chl-a tend to be more accurate than those for chl-b from the plant images, in which the color of the leaves is close to blue-green. In our proposed prediction system, the yellow and the red pigments are considered as noises and removed from the extracted green pixels by the image processing.

Artificial Neural Network model (multilayer perceptron neural network)

The multilayer perceptron neural network (MLPNN) with a back propagation-learning algorithm was adopted to obtain the model for predicting St. Jon's wort leaf chlorophyll content. The MLPNN model was trained using the reflectance of St. John's wort leaves detected at RGB components as the inputs and the corresponding measured SPAD readings, Chl-a and Chl-b as the outputs. In this research, MLPNN model consisted of 3 neurons in the input layer (R, G, and B) and 3 neurons in the output layer (SPAD readings, Chl-a and Chl-b) to match the training data set.

In the present study, 70% of the data set (1050 images) was selected as training data. The rest of the data set were used 15% for testing and 15% for validation predetermined values for the output error (MSE) and maximum iteration number were set to 0.001 and 1000 epoch, respectively. Since the accuracy of estimation is highly dependent on covering all level of data, the randomization process was repeated until a satisfactory level of data distribution was reached. The error back propagation (BP) training algorithm compares the estimated output value obtained from the model with the corresponding measured value.

The training process will be completed when all weighing indices are fixed and the ANN model can accurately estimate the output data as a function of input values (Kawashima and Nakatani, 1998). The number of hidden layers and the number of neurons in hidden layer were determined by the training error method. Mean square error of prediction (MSE) and coefficient of determination (R^2) were considered to evaluate the performance of ANN model constructed here. These criteria were calculated using the following equations of 5 and 6:

$$MSE = \frac{\sum_{i=1}^{N}\left(y_{ai} - y_{pi}\right)^2}{N-1} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \text{(eq. 3)}$$

$$R^2 = \frac{\left(\sum_{i=1}^{N}\left(y_{ai} - \bar{y}_a\right)\left(y_{pi} - \bar{y}_p\right)\right)}{\sum_{i=1}^{N}\left(y_{ai} - \bar{y}_a\right)^2\left(y_{pi} - \bar{y}_p\right)^2} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \text{(eq. 4)}$$

Where in the above equations y_i and y_{pi} are actual and predicted leaf SPAD values respectively. Different number of hidden layers, different number of neurons in each hidden layer, and different transfer functions between layers were taken into consideration. The optimum model obtained from this examination consisted of 1 hidden layer with 10 neurons and the sigmoid function in the hidden layer and the linear function in the output layer. To validate the model, the 30% of data set (450 images) was used as the input to the model (Figure 2).

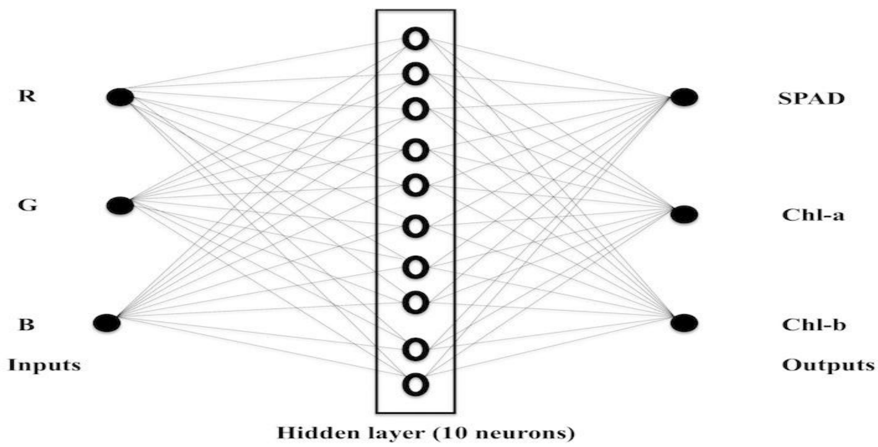
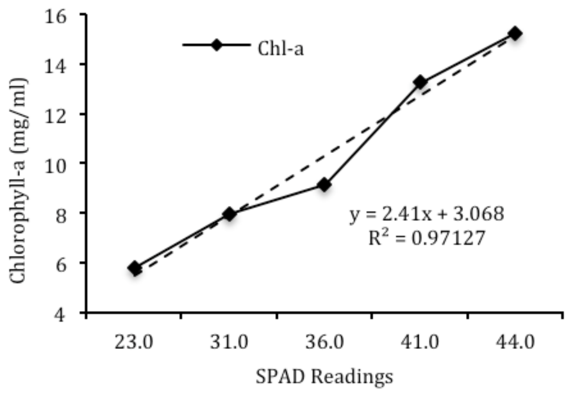


Figure 2. ANN Model: inputs are red, green, and blue components (RGB). Outputs are Chl-a, Chl-b and SPAD.

The hidden layer provided the ability of the network to generalize. Selecting appropriate number of hidden nodes was very important for building an ANN model. Considering the accuracy and the speed of convergence were two critical parameters for the network.

Result and Discussion

The linear relationships between SPAD readings and leaf Chlorophyll (Chl) content across all sampling were found in this study. The coefficient of determination (R^2) represents the dispersion of the points from the best-fit regression line. Thus, R^2 was used to define the best-fit relationship between Chl and SPAD readings estimation (Figure 3).



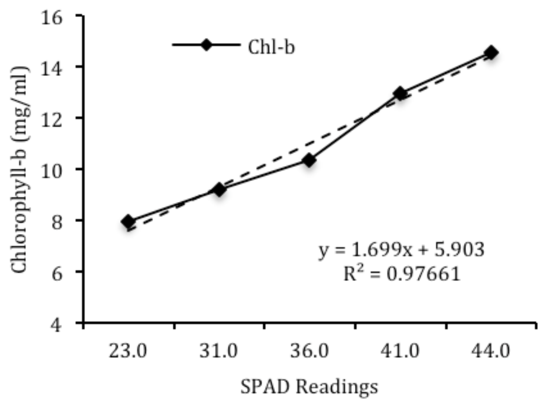


Figure 3. The relationships between Chlorophyll-a (mg/ml) and SPAD readings, Chlorophyll-b (mg/ml) and SPAD readings.

The equations of chl-a (y=2.34x + 2.138) and chl-b (y=1.714x + 6.152) explained 97.1% for chl-a, 97.6% and for chl-b of variation in the chlorophyll content of SPAD readings (Figure 1). The high R^2 values showed that chlorophyll meters are an effective tool for estimating chlorophyll concentration in plants. For this reason, SPAD readings data were used in ANN model.

ANN model

The results indicated that the ANN model trained with the RGB component was capable of the prediction of chlorophyll level of St. John's wort leaf (R^2 = 0.99). The results also showed that the image processing is a suitable approach for searching of plant chlorophyll level.

The inputs (R, G, and B components) of ANN model have been proven that leaf chlorophyll contents are in high correlation with RGB components in plant images.

The number of the neurons in this system was determined to be 10 through the learning process seeking a balance between the RMSE and the training time.

The study of the obtained results by ANN model indicated remarkably higher accuracy in comparing the predicted and measured chlorophyll levels (MSE = 0.005, R^2 =0.99). The overall average relative errors of chl-a, chl-b and SPAD readings obtained from the proposed predictor against those from the experimental measurement were 0.86 % for chl-a, 1.31 % for chl-b, and 0.65 % for SPAD readings. It is strong evident that the accuracies of prediction in chlorophyll levels and SPAD readings by the prediction model of ANN-based are acceptable (Table 2).

The preference for the ANN model can also be attributed to the relationship between RGB components and leaf chlorophyll level, which was perfectly linear. The overall results obtained from the present experiment verified that the MLPNN model could provide accurate estimations of St. John's wort leaf chlorophyll.

Table 2. Comparison of St. John's wort leaf chlorophyll contents by the ANN-based predictor and the experimental measurement using UV-spectrometer.

Rates of SHS	Chl-a	Chl-a Relative error (%)	Chl-b	Chl-b Relative error (%)	SPAD	SPAD Relative error (%)
0% of SHS	5.79 (5.83)	0.69	7.84 (7.96)	1.51	23.12 (23.34)	0.94
0.25% of SHS	7.89 (7.99)	1.25	9.09 (9.18)	0.98	31.18 (31.35)	0.54
0.50% of SHS	9.09 (9.17)	0.87	10.16 (10.33)	1.65	36.23 (36.42)	0.52
SHS	12.88 (12.97)	0.69	13.07 (13.23)	1.21	41.07 (41.25)	0.44
200% of SHS	14.44 (14.56)	0.82	15.07 (15.25)	1.18	44.34 (44.69)	0.78
Overall mean		0.86		1.31		0.65

SHS: Standard Hougland Solution; Chl-a, chlorophyll a; Chl-b, chlorophyll b; SPAD, soil plant analysis development. The up values are from the ANN-based predictor and the down values in the parentheses are from the experimental measurement. Relative error (%) =$|(V_{true} – V_{measured})/V_{true}|$ x 100 (if V_{true} is not zero).

The general purpose of multiple linear regressions is to learn more about the relationship between several independent or predictor variables and a dependent or criterion variable. A linear regression model assumes that the relationship between the dependent variable and independent variable(s) is linear. ANN is non-linear data driven self-adaptive systems, and they can identify and learn correlated patterns between input data sets and corresponding target values, even when the underlying data relationship is unknown. Artificial neural networks have been widely used to model complex and nonlinear processes and systems.

Conclusion

This study demonstrated that the ANN could use to predict the chlorophyll level of St. John's wort by analyzing the leaf images using image processing. When we compared this method with other remote sensing approaches, it is low in cost and easy to perform. The research also developed an ANN model, trained with RGB components, which resulted in estimating SPAD readings and the chlorophyll level of the St. John's wort leaves with higher accuracy. Compared to traditional chlorophyll measurement, this prediction system provides an effective and automatic measurement of chlorophyll to monitor the plant growth status in real time. In addition, based on the strong correlation between the chlorophyll and the SPAD readings. Moreover, the proposed method can be extended to predict the SPAD readings and chlorophyll levels (chl-a and chl-b) for other field crops.

Acknowledgements

This research was supported by Scientific and Technological Research Council of Turkey (TUBITAK) and it was carried out of the Department of Agricultural & Biosystems Engineering, North Dakota State University, United States.

References

Chappelle EW, Kim MS, McMurtrey JE (1992). Ratio analysis of reflectance spectra (RARS): an algorithm for the remote estimation of the concentrations of chlorophyll A, chlorophyll B, and carotenoids in soybean leaves. Remote Sensing of Environment. 39: 239-247.

Chen CT, Chen S, Hsieh KW, Yang HC, Hsiao S, Yang C (2007). Estimation of leaf nitrogen content using artificial neural network with cross-learning scheme and significant wavelengths. Transactions of the ASAE. 50: 295-301.

Cirak C, Ayan AK, Odabas MS (2011). Seed Germination of Hypericum triquetrifolium and Hypericum heterophyllum L. Medicinal and Aromatic Plant Science and Biotechnology. 5 (11): 105-107.

Demuth H, Beale M (2000). Neural network toolbox, Version 4., The MathWorks, Inc. Natick, MA, USA.

Gautam RK, Panigrahi S (2007). Leaf nitrogen determination of corn plant using aerial images and artificial neural networks. Canadian Biosystems Engineering. 49: 71-79.

Kawashima S, Nakatani M (1998). An algorithm for estimating chlorophyll content in leaves using a video camera. Annals of Botany. 81: 49-54.

Koumpouros Y, Mahaman BD, Maliappis M, Passam HC, Sideridis AB, Zorkadis V (2004). Image processing for distance diagnosis in pest management. Computers and Electronics in Agriculture. 44: 121-131.

Markwell J, Osterman JC, Mitchell JL (2005). Calibration of the minolta SPAD-502 leaf chlorophyll meter. Photosynthesis Research. 16: 467-472.

Namrata J, Ray SS, Singh JP, Panigrahy S (2007). Use of hyperspectral data to assess the effects of different nitrogen applications on a potato crop. Precision Agriculture. 8: 225-239.

Noh H, Zhang Q, Shin B, Han S, Feng L (2006). A neural network model of maize crop nitrogen stress assessment for a multi- spectral imaging sensor. Biosystems Engineering. 94: 477-485.

Odabas MS, Temizel KE, Caliskan O, Senyer N, Kayhan G, Ergun E (2014). Determination of reflectance values of hypericum's leaves under stress conditions using adaptive network based fuzzy inference system. Neural Network World. 24 (1): 79-87.

Odabas MS, Leelaruban N, Simsek H, Padmanabhan G (2014). Quantifying impact of droughts on barley yield in north dakota, usa using multiple linear regression and artificial neural network. Neural Network World. 24 (4): 343-355.

Pagola M, Ruben O, Ignacio I, Humberto B, Edurne B, Pedro AT, Carmen L, Berta L (2009). New method to assess barley nitrogen nutrition status based on image color analysis comparison with SPAD-502. Computers and Electronics in Agriculture. 65: 213- 218.

Pydipati R, Burks TF, Lee WS (2006). Identification of citrus disease using color texture features and discriminate analysis. Computers and Electronics in Agriculture. 52: 49-59.

Temizel KE, Odabas MS, Senyer N, Kayhan G, Bajwa S, Caliskan O, Ergun E (2014). Comparision of some models for estimation of reflectance of hypericum leaves under stress conditions. Central European Journal of Biology. 9(12): 1226-1234.

Wang Y, Wang F, Huang J, Wang X, Liu Z (2009). Validation of artificial neural network techniques in the estimation of nitrogen concentration in rape using canopy hyperspectral reflectance data. International Journal of Remote Sensing. 30: 4493-4505.

Wellburn AR (1994). The spectral determination of chlorophylls a and b, as well as total carotenoids, using various solvents with spectrophotometers of different resolution. Journal Plant Physiology. 144: 307–313.

Woebbecke DM, Meyer GE, Bargen KV, Mortensen DA (1995). Color indices for weed identification under various soil, residue and lighting conditions. Transactions of the ASAE. 38: 259-269.