Identification of Nitrogen Content of *Vernonia amygdalina* Leave Based on Artificial Neural Network Modeling

Sandra [1], Retno Damayanti [1*], Mochamad Bagus Hermanto [1], Rut Januar Nainggolan [1], Danuh Kanara Anta[1], Arini Robbil Izzati[1], Siska Ratna Anggraeni[1], Mitha Saadiyah[1]

[1] Department of Biosystem Engineering, Faculty of Agricultural Technology, Universitas Brawijaya, Jl Veteran, Malang, Jawa Timur, 65145 Indonesia
damayanti@ub.ac.id

Abstract. The level of greenness or the content of chlorophyll in the leaves is one indicator of plant health, where plants that are fertile and have enough nutrients will look green on their leaves. This indicates that the nitrogen (N) content, which is the constituent of leaf chlorophyll, is fulfilled properly and increases plant productivity higher. Knowing the nitrogen content in a plant can inform nutritional needs and monitor plant development quickly and precisely. This research aims to develop a mathematical model to predict the chlorophyll and nitrogen content in leaves using a machine vision method with texture and color analysis. Texture analysis uses the color features of Grey, RGB, HSL, HSV, and L*a*b* and the color co-occurrence matrix (CCM). The best 8 features have been obtained using Correlation as a selection attribute. The best ANN model was selected from 75% of training data and 25% of validation data with a topology structure of 8-30-40-2 with a learning rate value of 0.1 and momentum 0.5, trainlm as the selected learning function, tansig the activation function in the hidden layer and output layer. The selected ANN structure produces a validation correlation coefficient (R) of 0.99073 and a validation MSE of 0.0793.

Keywords: Artificial Neural Network; Chlorophyll; Color; Nitrogen.

1 Introduction

As an herbal plant, *Vernonia amygdalina* has various benefits due to its pharmacological properties such as antioxidant, hepatoprotective, antibacterial, anti-inflammatory, anticancer, antiviral, anesthetic, antifungal, antiprotozoal, antimicrobial, antileukemia, analgesic and so on. This is because *Vernonia amygdalina* is rich in nutrients and phytochemicals [1-8]. The results of previous studies stated that the difference in chlorophyll levels and the level of leaf development, affect the content of a plant [9-10]. Nitrogen (N) is one of the nutrients that are very important for plant growth and is the main component of chlorophyll and protein closely related to leaf color, plant growth status, and yields [11]. The quantitative assessment of the nitrogen content of the leaves can inform nutritional needs and monitor plant development quickly, precisely, and

Q. D. Utama et al. (eds.), *Proceedings of the 7th International Conference on Food, Agriculture, and Natural Resources (IC-FANRES 2022)*, Advances in Biological Sciences Research 35,
https://doi.org/10.2991/978-94-6463-274-3_17

very helpful for the accurate processing of nitrogen fertilizers, which can increase the yield and quality of the crop [12].

Laboratory measurements carried out destructively require leaf extract, thus causing damage to the sample, takes time and cannot be used in further measurements [13]. The chlorophyll meter (SPAD-502), a simple, fast, and non-destructive method, has been widely used to estimate the leaf nitrogen content correlated with chlorophyll content [14-17]. Apart from the high price of the SPAD-502, the measurement area is small (6 mm^2), it takes a lot of repetition to get the appropriate results, and SPAD-502 measurements are often experience difficulty in distinguishing chlorophyll levels in plant ethics close to or above the optimal N supply [18-19].

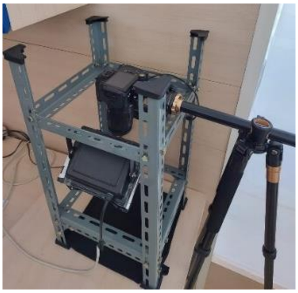
Recently, artificial intelligence technology has become represented by deep learning and become a hot spot of modern technology in agricultural research. An artificial neural network (ANN) is one of the computational techniques related to artificial intelligence, which works exactly with human neural systems. Each layer receives inputs with weights connected from other neurons, passing through neurons and generating output signals that other neurons can also generate. In this way, the process runs along the neurons and layers back and forth. When it reaches the specified error value, the network training process will stop and form a model. Each network is developed with algorithms and several other performance parameters whose performance results in different variations. Through the trial and error method, the most optimal results can be obtained [20-22].

In this study, an information-based ANN model developed through ten texture features from the Grey, RGB, HSL, HSV, and L*a*b* colors space used to identify nitrogen content in *Vernonia amygdalina*. The ANN model is built with a feed-forward back-propagation neural network architecture. In the end, the topology structure of the ANN model that has the best performance is obtained which is seen through the parameters of the lowest Mean squared error (MSE) value and a coefficient of determination (R) that is high or close to 1

2 Research Method

This study used *Vernonia amygdalina* leaf that grows wild in the Malang area, East Java, Indonesia. The materials used for nitrogen measurement were K2SO4, CuSO4, H3BO3, NaOH, H2SO4, KCL, aqueous, BCG-MR indicator, ice cubes, as well as boiling stones. The tool used for the acquisition of leaf imagery is a series of black boxes made of perforated L iron (Fig.1) and conditioned in a dark state covered with black cloth. On the Black Box is a Canon EOS 700D camera connected to the laptop. SPAD-502 Plus is used to measure the chlorophyll value or the level of leaf greenness. The software used is windows 10 64-bit operating system, visual basic 6.0 for color-based image feature extract and texture analysis, Waikato Environment for Knowledge Analysis (WEKA) 3.9.6 used for feature selection, MatlabR2021a as a framework for developing ANN models and Microsoft Excel 2021. The tools used for nitrogen measurement are a spatula, mortar, pestle, flask, kjeldahl, watch glass, electric stove, a series of distillation tools (heating mantle, condenser, pump, bucket hose), analytical balance sheet, beaker glass, measuring cup, drip pipette, fume hood, Erlenmeyer, funnel, and burette.

Vernonia amygdalina was picked at 07.00 WIB to 08.00 WIB using plant scissors. The *Vernonia amygdalina* used as samples in this study were 3–4 most-ended leaves for young leaves, 3–4 base leaves that were included in the category of old leaves, and leaves that grew in the middle for mature leaves (between young and old leaves). At each development, leaves are plucked by 100 strands. The leaf imagery was taken through a digital camera assembled on a tripod on a modified set of tools and covered with a black cloth to prevent outside light from entering. The images obtained are in the form of Bitmap format and each image is cropped with a size of 600x600 pixels. Before the *Vernonia amygdalina* image feature is extracted, an augmentation treatment is given to add to the variety of cloudy photos to improve the learning model in artificial neural networks. Each leaf image was rotated at an angle of 0º, 90º, 180º, and 270º.



(a)

Fig 1. Image acquisition set-up

In this study, the leaf image was transformed from the RGB color space to the HSV, HSL, L*a*b*, and gray systems. Then, a matrix co-occurrence is made for each color space, namely grey-GLCM, *red* CCM, green-CCM, blue-CCM, hue CCM, saturation (HSL) CCM, saturation (HSV) CCM, value CCM, lightness CCM, L* CCM, a* CCM, b* CCM. The final stage of this process is to use ten Haralick equations to calculate texture features as texture analysis. The ten texture features are entropy, energy, contrast, homogeneity, sum mean, variance, correlation, maximum probability, inverse difference moment, and cluster tendency. The value of the texture feature is obtained from the following calculations [23]:

$$Energy = \sum_i^M \sum_j^N P^2\,[i,j] \tag{1}$$

$$Entropy = -\sum_I^M \sum_j^N P\,[i,j]\,log\,P[i,j] \tag{2}$$

$$Contrast = \sum_i^M \sum_j^N (i-j)^2\,P\,[i,j] \tag{3}$$

$$Homogeneity = \sum_i^M \sum_j^N \frac{P[i,j]}{1+|i-j|} \tag{4}$$

$$Inverse\ Difference\ Moment = \sum_i^M \sum_j^N \frac{P[i,j]}{|i-j|^k} \tag{5}$$

$$Correlation = \sum_i^M \sum_j^N \frac{(i-\mu)(j-\mu)P[i,j]}{\sigma^2} \tag{6}$$

$$Sum\ Mean = \frac{1}{2}\sum_i^M \sum_j^N (iP[i,j] + jP[i,j]) \tag{7}$$

$$Variance = \frac{1}{2}\sum_i^M \sum_j^N ((i-\mu)^2 P[i,j] + (j-\mu)^2 P[i,j]) \tag{8}$$

$$Cluster\ Tendency = \sum_i^M \sum_j^N (i+j-2\mu)^k P\,[i,j] \tag{9}$$

$$Maximum\ Probability = Max_{i,j}^{M,N} P[i,j] \tag{10}$$

where P (I,j) is the $(I,j)^{th}$ element of the normalized matrix of shared events, the μ and σ are the mean and standard deviation of the pixel element. All ten texture features were extracted at a distance (d=1) and an angle (θ = 0).

The chlorophyll content or the level of *Vernonia amygdalina* greenness was measured using the chlorophyll analyzer SPAD-502 Plus. For leaves that have a small size (usually young leaves), five to ten SPAD values are taken while for other leaf sizes 15–20 SPAD values are taken. Then the average value of each measurement becomes the chlorophyll value or the level of leaf greenness. The measurement of nitrogen content of the leaves was measured using the Kjeldahl method. The measurement values of SPAD values and leaf nitrogen content are the dataset that will be predicted using ANN modeling [15, 24].

A total of 1200 images is divided into 75% as training data and 25% as validation data. Then, all data obtained were normalized in the ranges of -1 and 1, before conducting ANN modeling. In designing the best ANN topology structure, involves through sensitivity analysis that will evaluate how the parameters in the output layer respond to variations contained in the model-building components, such as learning functions, activation functions, learning rate (lr), and momentum (mc) (0.1, 0.5, 0.9), the number of hidden layers (1;2) and the number of nodes (10, 20, 30, 40). There are 3 activation functions described in this study, namely, purelin, tansig, and logsig. The selection of the best performance is based on the lowest MSE (*error*) value and the highest determinant coefficient (R) value.

3 Result and Discussion

Fig. 2 shows the results of the digital image acquisition of *Vernonia amygdalina* leaves from various variations in leaf age. There are 8 texture and color features were selected, namely Light (LAB) Entropy, b*(LAB) Entropy, Green Correlation, Red Maximum Probability, Value Energy, Gray Correlation, Green Maximum Probability, and Red Correlation, which becomes data input that has a relationship with the nitrogen content of the leaves. The results obtained were used as datasets to build and train ANN models to identify nitrogen content in *Vernonia amygdalina*.

(a) (b) (c)

Fig 2. Example of *Vernonia amygdalina* leaves: (a) young leave; (b) mature leave; (c) old leave

Table 1. Trial and error learning function

Learning Function	R Training	R Validation	MSE Training	MSE Validation
Traincgb (*Conjugate Gradient Backpropagation with Powell-Beale Restarts*)	0.98204	0.98791	0.0100	0.0937
Traincgf (*Conjugate Gradient Backpropagation with Fletcher-Reeves Updates*)	0.97296	0.98689	0.0150	0.1068
Traincgp (*Conjugate Gradient Backpropagation with Polak Ribiere Update*)	0.97692	0.9888	0.0127	0.0956
Traingd (*Gradient Descent Backpropagation*)	0.75573	0.98008	0.1198	0.1560
Traingda (*Gradient Descent with Adaptive Learning Rate Backpropagation*)	0.74262	0.97875	0.1254	0.1519
Traingdm (*Gradient Descent with Momentum Backpropagation)*	0.76292	0.97987	0.1167	0.1490
Traingdx (*Gradient Descent with Momentum and Adaptive Learning Rate Backpropagation*)	0.7512	0.97916	0.1217	0.1452
Trainlm (*Lavenberg-Marquadt Backpropagation)*	**0.98194**	**0.99073**	**0.0100**	**0.0793**
Trainoss (*One-Step Secant Backpropagation*)	0.9754	0.9889	0.0136	0.0876
Trainrp (*Resilient Backpropagation*)	0.96497	0.98664	0.0192	0.0990
Trainscg (*Scaled Conjugate Gradient Backpropagation*)	0.98199	0.98617	0.0100	0.1048

Table 1 shows the performance of the ANN models developed with 11 different learning functions. The lowest validation MSE value is generated using the Lavenberg-Marquadt Backpropagation learning function or better known as the trainlm learning function. The trainlm algorithm is one of the Gauss-Newton-type method with optimization techniques that utilize the lowest drops for complex non-linear patterns [25]. In this study, the MSE of validation was 0.0793 and the value of the coefficient of determination was close to 1, namely, 0.99073. The tansig activation function was selected as a non-linear activation function on the first, second hidden layer and output layer by generating the lowest validation MSE value of 0.0793 and an R value of 0.99073 (Table 2). The tansig activation function is a hyperbolic tangent transfer function associated with bipolar sigmoids that have an output value in the range between -1 to +1 [26]. In Table 3, the lowest validation MSE value of 0.0793 with two hidden layers. Hidden layers are other layers besides the output layer [27]. Choosing the right number of hidden layers must be performed to avoid overfitting and underfitting conditions. Overfitting is a condition that occurs when a hidden layer has an enormous number compared to the complexity of the problem. An increase in the number of hidden layers will result in a gradual improvement in the performance of Back-propagation [28]. The selected ANN model is constructed with an 8-30-40-2 topology structure, which can be seen in Fig. 3.

Table 2. Trial and error activation function

Learning Function	Activation Function			R Training	R Validation	MSE Training	MSE Validation
	Hidden Layer 1	Hidden Layer 2	Output Layer				
Trainlm	Tansig	Tansig	Purelin	0.98224	0.98512	0.0098	0.1425
	Tansig	**Tansig**	**Tansig**	**0.98194**	**0.99073**	**0.0100**	**0.0793**
	Tansig	Tansig	Logsig	0.81581	0.97249	0.1884	0.2693
	Logsig	Logsig	Purelin	0.98233	0.98837	0.0098	0.1136
	Logsig	Logsig	Tansig	0.98246	0.98727	0.0097	0.1078
	Logsig	Logsig	Logsig	0.84622	0.97292	0.1832	0.2556

Table 3. Trial and error ANN topology

Learning Rate	**Momentum**	**ANN Topology**	**R Training**	**R Validation**	**MSE Training**	**MSE Validation**
0.1	**0.5**	8-30-2	0.98202	0.984	0.0100	0.1312
		8-40-2	0.9823	0.98449	0.0099	0.1171
		8-30-30-2	0.98217	0.9852	0.0099	0.1129
		8-30-40-2	**0.98194**	**0.99073**	**0.0100**	**0.0793**
		8-40-40-2	0.98246	0.98841	0.0097	0.0875
	0.9	8-30-2	0.98197	0.97992	0.0100	0.1492
		8-40-2	0.98206	0.98344	0.0100	0.1215
		8-30-30-2	0.9821	0.98903	0.0099	0.0816
		8-30-40-2	0.98245	0.98799	0.0097	0.0950
		8-40-40-2	0.98207	0.98738	0.0100	0.1029

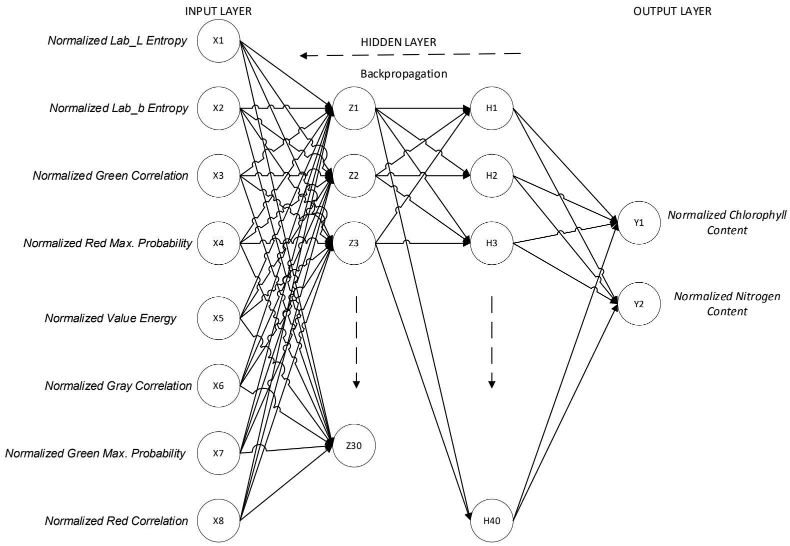
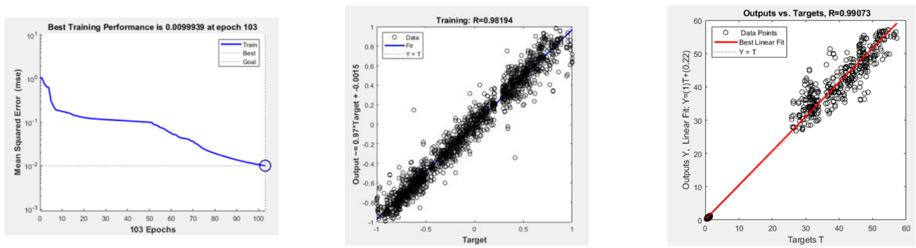


Fig 3. ANN Model



(a) (b) (c)

Fig 4. Performance of ANN model: (a) Back-propagation Neural Network learning process on training data set; (b) Result regression plot for training; (c) Result in regression plot for validation

Based on the squared error mean shown in Fig. 4(a), the best training performance was achieved at the 103rd epoch, with an MSE parameter performance value of 0.0099939. In the development of this ANN model, the maximum number of epochs is 10000 with a goal of 0.01, meaning that training will stop after 10000 epochs or when the goal value of 0.01 has been achieved. Fig. 4(a) shows the regression plot on training simulation and validation simulation Fig. 4(b). The distribution of data getting closer to the linear line of conformity indicates that the prediction is getting closer to the actual value. As the R-value in the simulation validation data is close to 1, which is 0.99073, this shows that the model has good performance. Based on the results of this study, an R-value that close to 1, shows that there is a relationship between the data on the texture characteristics of the color space in the *Vernonia amygdalina* leaf image, namely, Light (LAB) Entropy, b*(LAB) Entropy, Green Correlation, Red Maximum Probability, Value Energy, Gray Correlation, Green Maximum Probability, and Red Correlation to SPAD values with leaf nitrogen content of *Vernonia amygdalina*.

4 Result and Discussion

The best ANN model to predicting of nitrogen content using the 8-30-40-2 topology (8 inputs, 30 nodes in the first hidden layer, 40 nodes in the 2nd hidden layer, 2 outputs) with a learning rate parameter of 0.1 and a momentum of 0.5, learning functions trainlm, the function of activating tension on the hidden layer and output layer. The selected ANN model generates an MSE training value of 0.0100; MSE validation 0.0793 with a value of R training 0.98194; Validation R value of 0.99073. These results indicate that digital image processing with the development of the ANN model has the potential as a sensor to detect nitrogen content in *Vernonia amygdalina*.

References

1. Alara, O.R., Abdurahman, H., Kholijah, S., Mudalip, A., Olalere, A. : Phytocemical and Pharmachological Properties of Vernonia amygdalina: A Review. Journal Chemical Engineering and Industrial Biotechnology (JCEIB) Open Access 2, 80-96 (2017).
2. Alara, O.R., Abdurahman, N.H., Olalere, O.A. : Ethanolic extraction of flavonoids, phenolics and antioxidants from Vernonia amygdalina leaf using two-level factorial design. Journal of King Saud University - Science, 32(1), 7–16 (2020).
3. Ijeh, I. I., Ejike, C. E. C. C. : Current perspectives on the medicinal potentials of Vernonia amygdalina Del. Journal of Medicinal Plants Research, *5*(7), 1051–1061 (2011).
4. Kaur, D., Khalsa, G. H. G., Kaur, N., & Chopra, A. : A comprehensive review on phytochemistry and pharmacological activities of Vernonia amygdalina. Journal of Pharmacognosy and Phytochemistry, 8(3), 2629–2636 (2019).
5. Nursuhaili, A. B., Nur Afiqah Syahirah, P., Martini, M. Y., Azizah, M., Mahmud, T. M. M.: A review: Medicinal values, agronomic practices and postharvest handlings of vernonia amygdalina. Food Research, 3(5), 380–390 (2019).
6. Owoeye, O., Yousuf, S., Akhtar, M. N., Qamar, K., Dar, A., Farombi, E. O., Onwuka, S. K., Choudhary, M. I. : Another anticancer elemanolide from Vernonia amygdalina Del. International Journal of Biological and Chemical Sciences, 4(1), 226–234 (2010).
7. Quasie, O., Zhang, Y. M., Zhang, H. J., Luo, J., Kong, L. Y. : Four new steroid saponins with highly oxidized side chains from the leaves of Vernonia amygdalina. Phytochemistry Letters, 15, 16–20. (2016).
8. Yeap, S. K., Ho, W. Y., Ky, H., Alitheen, N. B. : Vernonia amygdalina, an ethnoveterinary and ethnomedical used green vegetable with multiple bioactivities. Journal of Medicinal Plants Research, 4(25), 2787–2812 (2010).
9. Abdulkadir, A. R., Jahan, M. S., Zawawi, D. D. : Effect of chlorophyll content and maturity on total phenolic, total flavonoid contents and antioxidant activity of Moringa oleifera leaf (Miracle tree). Journal of Chemical and Pharmaceutical Research, *7*(5), 1147–7384 (2015).
10. Sreelatha, S., Padma, P. R. : Antioxidant activity and total phenolic content of Moringa oleifera leaves in two stages of maturity. Plant Foods for Human Nutrition, 64(4), 303–311 (2009).

11. Fageria, N. K., Baligar, V. C., Jones, C. A. : Growth and Mineral Nutrition of Field Crops. 3rd edition. CRC Presss, Taylor & Francis Group, New York (2011).
12. Yang, B., Zhu, Y., Zhou, S. : Accurate wheat lodging extraction from multi-channel UAV images using a lightweight network model. Sensors, 21(20), 1–16 (2021).
13. Uddling, J., Gelang-Alfredsson, J., Piikki, K., Pleijel, H. : Evaluating the relationship between leaf chlorophyll concentration and SPAD-502 chlorophyll meter readings. Photosynthesis Research, 91(1), 37–46 (2007).
14. Ata-Ul-Karim, S. T., Cao, Q., Zhu, Y., Tang, L., Rehmani, M. I. A., Cao, W. : Non-destructive assessment of plant nitrogen parameters using leaf chlorophyll measurements in rice. Frontiers in Plant Science, 7 (2016).
15. Garcia-Servin, M., Contreras-Medina, L. M., Torres-Pacheco, I., Guevara-González, R. G.: Estimation of nitrogen status in plants. Nitrate Handbook: Environmental, Agricultural, and Health Effects, CRC Press, Taylor & Francis Group, New York (2021).
16. Valverde, J.C. : Estimation of leaf nitrogen content from non-destructive methods in eucalyptus tereticornis and eucalyptus saligna plantations. Revista Facultad Nacional de Agronomia Medellin, 74(3), 9655–9666 (2021).
17. Wang, Y., Wang, D., Shi, P., Omasa, K. : Estimating rice chlorophyll content and leaf nitrogen concentration with a digital still color camera under natural light. Plant Methods, 10(1) (2014).
18. Purcell, L. C., Mozaffari, M., Karcher, D. E., Andy King, C., Marsh, M. C., Longer, D. E. : Association of "Greenness" in corn with yield and leaf Nitrogen concentration. Agronomy Journal, 103(2), 529–535 (2011).
19. Zhang, J., Blackmer, A. M., Ellsworth, J. W., Koehler, K. J. : Sensitivity of chlorophyll meters for diagnosing nitrogen deficiencies of corn in production agriculture. Agronomy Journal, 100(3), 543–550 (2008).
20. Sinshaw, T. A., Surbeck, C. Q., Yasarer, H., Najjar, Y. : Artificial Neural Network for Prediction of Total Nitrogen and Phosphorus in US Lakes. Journal of Environmental Engineering, 145(6), 1–11 (2019).
21. Tasdemir, S. : Artificial Neural Network Model for Prediction of Tool Tip Temperature and Analysis. International Journal of Intelligent Systems and Applications in Engineering, 6(1), 92–96 (2018).
22. Tümer, A.E., Koç, B.A., Koçer, S. Artificial Neural Network Models for Predicting the Energy Consumption of the Process of Crystallization Syrup in Konya Sugar Factory. International Journal of Intelligent Systems and Applications in Engineering, 5(1), 18–21 (2017).
23. Haralick, R. M., Dinstein, I., Shanmugam, K. : Textural Features for Image Classification. IEEE Transactions on Systems, Man and Cybernetics, SMC-3(6), 610–621 (1973).
24. Saberioon, M. M., Amin, M. S. M., Anuar, A. R., Gholizadeh, A., Wayayok, A., Khairunniza-Bejo, S. : Assessment of rice leaf chlorophyll content using visible bands at different growth stages at both the leaf and canopy scale. International Journal of Applied Earth Observation and Geoinformation, 32(1), 35–45 (2014).
25. Taghavifar, H., Mardani, A. : Application of artificial neural networks for the prediction of traction performance parameters. Journal of the Saudi Society of Agricultural Sciences, 13, 35-43 (2014).
26. Dorofki, M., Elshafie, A.H., Jaafar, O., Karim, O.A., Mastura, S. : Comparison of artificial neural network transfer functions abilities to simulate extreme runoff data. International Conference on Environment, Energy and Biotechnology, 33, 39-44 (2012).
27. Ogunbo, J.N., Alagbe, O.A., Oladapo, M.I., Shin, C. : N-hidden layer artificial neural network architecture computer code: geophysical application example. Heliyon 6(6), 1 (2020).

28. Asthana, S., Pandit, A., Bhardwaj, A. : Analysis of multiple hidden layer vs. accuracy in performance using back propagation neural network. Indian Journal of Science Technology, 10(4), 1-4 (2017).