Plant Growth Prediction Model of Red Chili (*Capsicum annuum L.*) by Different Manipulation Environment

G M D Putra[1,2], L Sutiarso[1,*] A P Nugroho[1] Ngadisih[1] M S I Chaer[1]

[1] *Department of Agricultural and Biosystems Engineering, Faculty of Agricultural Technology, Universitas Gadjah Mada, Yogyakarta, Indonesia*
[2] *Department of Agricultural Engineering, Faculty of Food and Agroindustrial Technology, Universitas Mataram, Mataram, Indonesia*
[*]*Corresponding author. Email: lilik-soetiarso@ugm.ac.id*

ABSTRACT

Several factors influence plant growth, including sun intensity, nutrient content, soil moisture, temperature, genes, and hormones. Many studies have been carried out in constructing plant growth models to simulate plant growth in different treatments. This study aims to develop a mathematical model with a linear regression approach and an artificial neural network. This research method used an experimental design using three treatments consisting of control (T1), 50% shade (T2), and 80% shade (T3). Each treatment had five replications of the chili plant. The tools and materials used were red chili (*Capsicum annuum L.*) seeds of 30 DAP, a greenhouse of 3 x 3 meters, a drip irrigation control system, 25 x 30 cm polybags, and fertile soil media. The results showed that linear regression models of the 1st and 2nd order could be used to predict plant growth with an average RMSE value of 1.53. In contrast, the use of artificial neural networks showed a smaller RMSE value of 0.12 which means that the artificial neural network method was better at predicting plant growth.

Keywords: *ANN, model, red chili, regression.*

1. INTRODUCTION

Red chili is a horticultural product that has an essential role in Indonesia. Red chili was once a contributor to inflation in Indonesia in 2018, which was from 0.0.7% to 0.2% in 2019. The chili trade still has fluctuating prices; under certain conditions, the price of chili is high, and on other occasions, the price drops [1,2].

Red chili production is determined by plant growth. Plant growth is influenced by internal (genes) and external (climate) factors [3]. Climatic conditions following chili plants' growth will produce optimal production. The previous researcher explained that tomato plant height was positively correlated with yield [4][5][6]. These studies are also supported by other studies that demonstrate that the increase in the height of red chili plants by 21.8% can improve production by 20.2% [7].

Plant growth data can be used to make a model to predict plant growth. Many previous studies have carried out models for predicting plant growth. The first researcher made a chili plant growth model using the linear regression method [8], the other developed a regression model to predict pepper plant growth [9], and the last created a lettuce growth model using Artificial Neural Network (ANN) [10]. Furthermore, the ANN and the regression model can also predict the Ajowan (*Carum copticum L.*) plant oil content [11].

In this study, the linear regression model and the ANN model were used to predict the growth of red chili plants. The best model was then applied to predict red chili plants' growth from these two models.

A. D. Saputro et al. (Eds.): ICOSEAT 2022, ABSR 26, pp. 19–27, 2023.
https://doi.org/10.2991/978-94-6463-086-2_4

2. MATERIALS AND METHODS

2.1. Experimental details

This research was carried out in the Banguntapan area, Bantul, Special Region of Yogyakarta, with the coordinates 7°50'23.1"S 110°22'47.5"E from February 2021 to April 2021. Chili plants were planted inside a 60 m^2 screen net greenhouse, as shown in Figure 1. Outside the greenhouse, a control treatment of the chili plant was placed.

Figure 1 The screen net greenhouse

The type of chili used in this research was Ta-Nvi curly chili seeds produced by Scani Seed Indonesia. The plant media used came from Citra Organic Gro, made by CV. Citra Berlian Indonesia. Chili plants were planted in polybags 25 x 30 cm. The irrigation system used a drip irrigation system consisting of a 3/4-inch pipe, 12-hole manifold with a diameter of 7 mm, HDPE (high-density polyethylene) hose with a diameter of 7 mm, manual valve, and an emitter with a diameter of 7 mm. The microclimate was observed in the field using a set of monitoring systems based on Arduino microcontrollers with temperature, humidity, and light intensity sensors. The monitoring system used sensors that have also been carried out by previous researchers [12][13]. Measuring plant height was conducted using a ruler.

The research design used 3 (three) types of treatment, *i.e.*, control, 50% of shade, and 80% of shade. The data parameters in this study were temperature, humidity, light intensity, and plant height. Microclimate data were collected every 5 minutes, then averaged in the morning, afternoon, and evening for each treatment until 95 days. Plant height was measured weekly using a random sample from three chili plants for each treatment. Figure 2 shows the research flow chart.

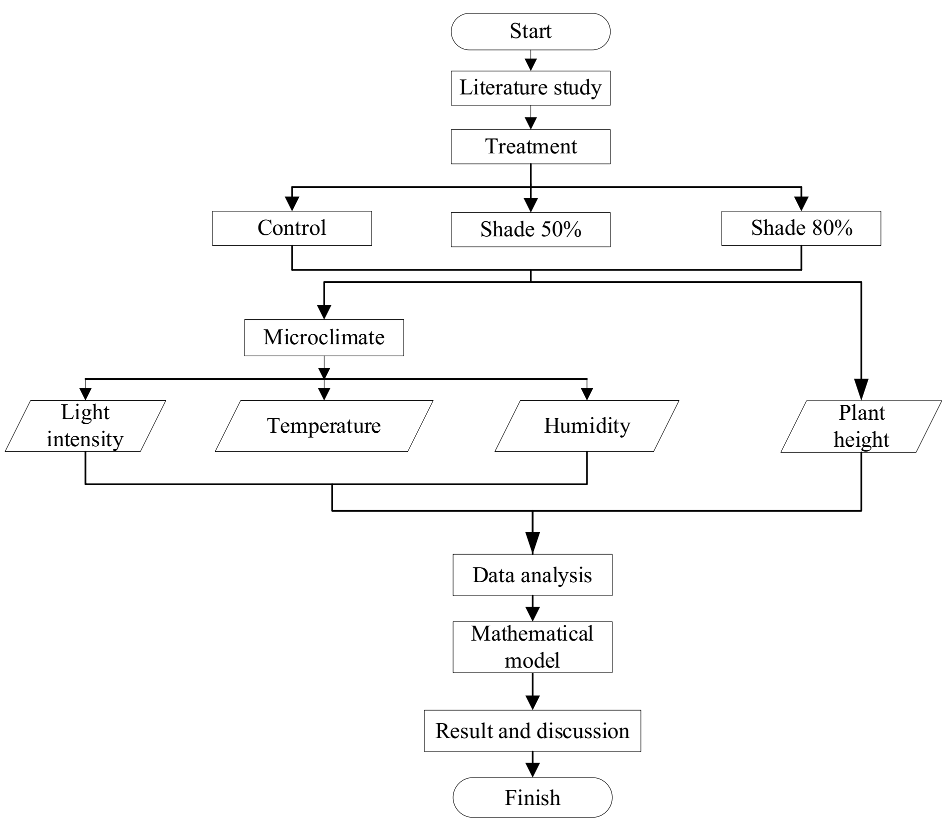


Figure 2 Experimental research flow chart

2.2. Mathematical model design

This research built a mathematical model based on plant growth. The regression equation model is a simple form and is one of the models commonly used in multivariate statistical analysis [14]. This model measured plant growth based on a close relationship between actual plant growth as the dependent variable and other components as independent variables. The suitable regression model measured plant growth accurately; therefore, finding the best regression model to determine plant growth was necessary.

The linear regression model ($y = a + bx$) has a positive correlation between the dependent variable (y) with the independent variable (x) for any available independent variables. However, some regression models do not always show a positive correlation. The 2nd order regression ($y = ax^2 + bx + c$) has a maximum point, so there will be a negative correlation after passing the maximum point. The 3rd order regression ($y = ax^3 + bx^2 + cx + d$) also has a turning point which causes the modeling to be biased to achieve its maximum point. Therefore, the multiplication of leaf length and width data in the 2nd and 3rd order regression modeling must be limited so that the correlation has no bias and is always positive [15,16,17]. Generally, the higher the order of the regression model, the better the model in data forecasting, as presented in previous research [18].

Besides using a regression model, this study used an Artificial Neural Network (ANN) model. ANN is defined as an information processing system with performance characteristics based on the modeling of biological neural systems through the approach of biological computation properties [19]. ANN is an alternative solution in discriminant analysis with non-linear functions or non-linear regression models [20]. The advantage of this method is that no assumption is needed where the data does not have to be normally distributed multivariate or does not require a model yet superior due to its high accuracy [21]. ANN consists of several neurons or nodes or processing elements, and there is a relationship between these neurons. Neurons transform the information (signals) received through the output path to other neurons; this relationship is known as weight [22]. Figure 3 is a simple structure of ANN with its activation function.

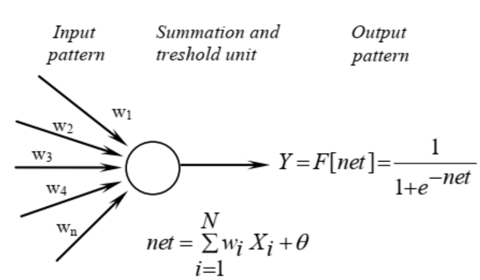


Figure 3 ANN structure

One of the ANN models used to solve the problem in this research was the Back Propagation Network. The backpropagation algorithm architecture consists of three layers, *i.e.*, the input layer, the hidden layer, and the output layer. There was no computing process at the input layer, but there was a transmission process of the input signal x to the hidden layer [23]. There was a computational process on the weights and biases in the hidden and output layers. The amount of output from the hidden and output layers was calculated based on certain activation functions. This backpropagation algorithm used a binary sigmoid activation function since the expected outcome was between 0 and 1. Figure 4 shows the backpropagation architecture.

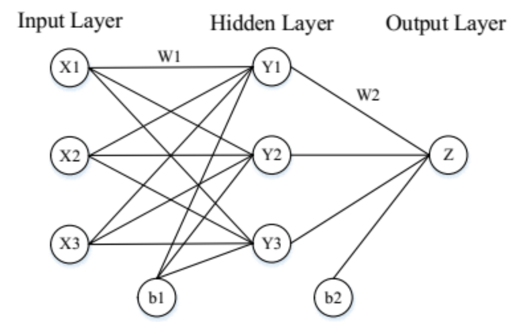


Figure 4 The backpropagation architecture

The performance of ANN-backpropagation can be affected by the number of iterations and the performance of the goal. Furthermore, the length of training before the termination condition is reached also determines its performance. To observe the network response, other data that is not executed to train the network is used, which is commonly called test data (validation data). The network can be trained continuously until the training data pattern is recognized perfectly, but this does not guarantee that the network will be able to identify the test data pattern correctly [23].

2.2.1. Accuracy and error testing

At this stage, the obtained prediction results were analyzed by observing the accuracy level and error in the system. Testing aimed to determine whether the input, process, and output work systems were in accordance with the expected goals. RMSE (Root Mean Square Error) was used to calculate the error value. The RMSE value indicates how much error in the prediction result is compared to the actual value. The RMSE was calculated using the absolute error in each period divided by the amount of data observed for the appropriate period. The smaller the RMSE value, the better the prediction performance [24]. The RMSE formula can be seen in equation 1.

$$RMSE = \sqrt{\frac{\sum(Aktual - Prediction)^2}{n}} \quad (1)$$

In creating the model, several stages were carried out, *i.e.*, data collection, initial processing, plant growth prediction, accuracy test, and implementation, as shown in Figure 5.

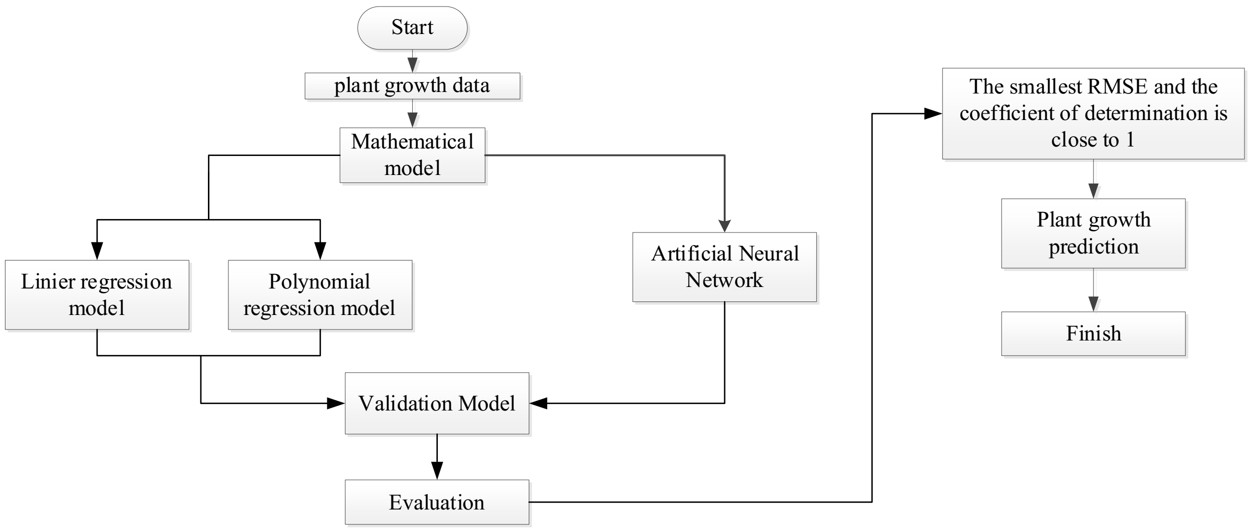


Figure 5 Design of red chili (*Capsicum annuum L.*) plant growth mathematics model

3. RESULTS AND DISCUSSION

The growth of red chili plants was measured in each treatment, *i.e.,* control, 50% shade, and 80% shade. The three treatments showed that plant growth was positively correlated with time. Figure 6 shows the growth of the red chili plant.

Plant growth in the control treatment was almost the same as in the 80% shade treatment. The control treatment had a plant height of 62.1 cm, the 50% shade treatment was 55.4 cm, and the 80% shade treatment was 59.6 cm. Plants in the 50% shade treatment had the lowest plant height compared to other treatments. In this condition, this indicated that the 50% treatment was not suitable for the red chili plants' growth, as well as in the 80% shade treatment, although they had almost the same plant height as the control treatment, this growth was not good because of the lack of light (etiolation) [25]. This phenomenon occurs due to the presence of non-green plastids (etioplasts) in plant tissues [26], as the light level is too low for chloroplast maturation [27].

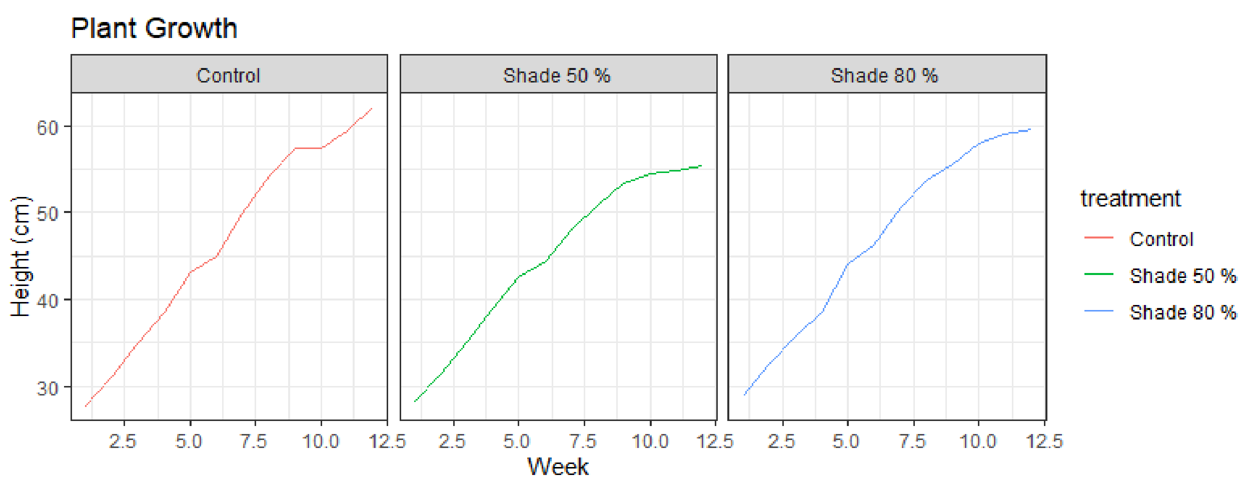


Figure 6 Data of plant height on different treatment

Chili plant growth data in each treatment was then constructed into a mathematical model using the regression and ANN models. The regression model used linear and polynomial regression of 2nd order, while the ANN model developed was simple, with only one input and one output.

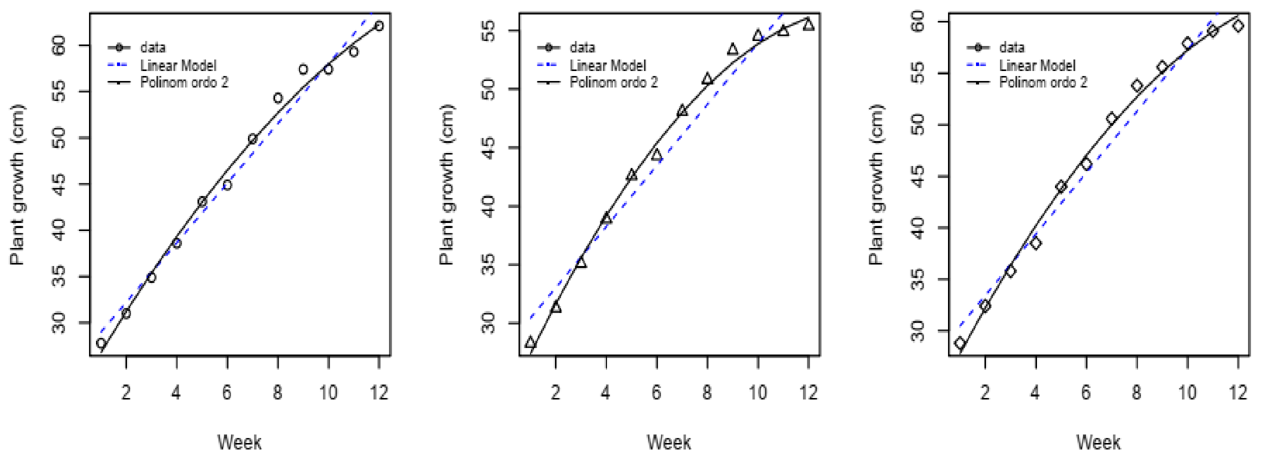


Figure 7 Plant growth model on different treatments using linear regression: (a) control, (b) 50% shade, (c) 80% shade

The comparison of the regression model between the actual and predicted values can be seen in Figure 7. These results show that the linear regression model for the control treatment, 50% shade treatment, and 80% shade treatment has a coefficient of determination (R^2) of 0.97, 0.95, and 0.97, respectively, and the second-order polynomial model gives the coefficient of determination (R^2) of 0.99, 0.99, 0.99. The value of the coefficient of determination closest to the value of 1 means it has a better model [28]. The regression model equation for each treatment can be shown in Table 1.

Table 1. The regression equation model for each treatment

No	Treatment	Regression equation			
		Linear Regression	RMSE	2nd order polynomial regression	RMSE
1	control	y=3.2241x+25.7682	2.83	$y=-0.11996x^2+4.78354x+22.12955$	1.73
2	50% shade	y=2.611x+27.821	3.74	$y=-0.16531x^2+4.75987x+22.80682$	1.28
3	80% shade	y=2.9829x+27.4697	3.25	$y=-0.14503x^2+4.86826x+23.07045$	1.60

Table 1 shows that the smallest RMSE value of the three treatments is the 2nd-order polynomial regression model. The smaller the RMSE value, the better the resulting model. However, the polynomial regression model shows a negative constant value, which means that the plant height will decrease or die at a specific time in accordance with the phytohormone that controls plant physiological action [29]. The relationship that describes the actual plant growth with the predicted plant growth is shown in Figure 8.

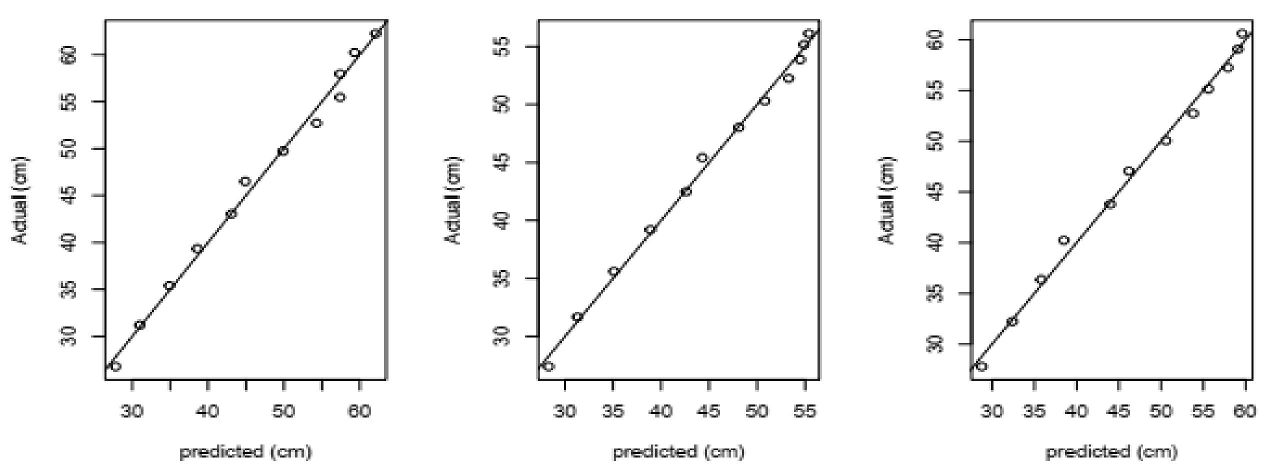


Figure 8 Actual and predicted plant growth using the linear regression model: (a) control, (b) 50% shade, (c) 80% shade

The architecture of red chili plant growth prediction using simple ANN can be seen in Table 2. This study used several variations of the hidden layer to get the best predictive value. The trial-and-error method was applied to determine the number of hidden layers.

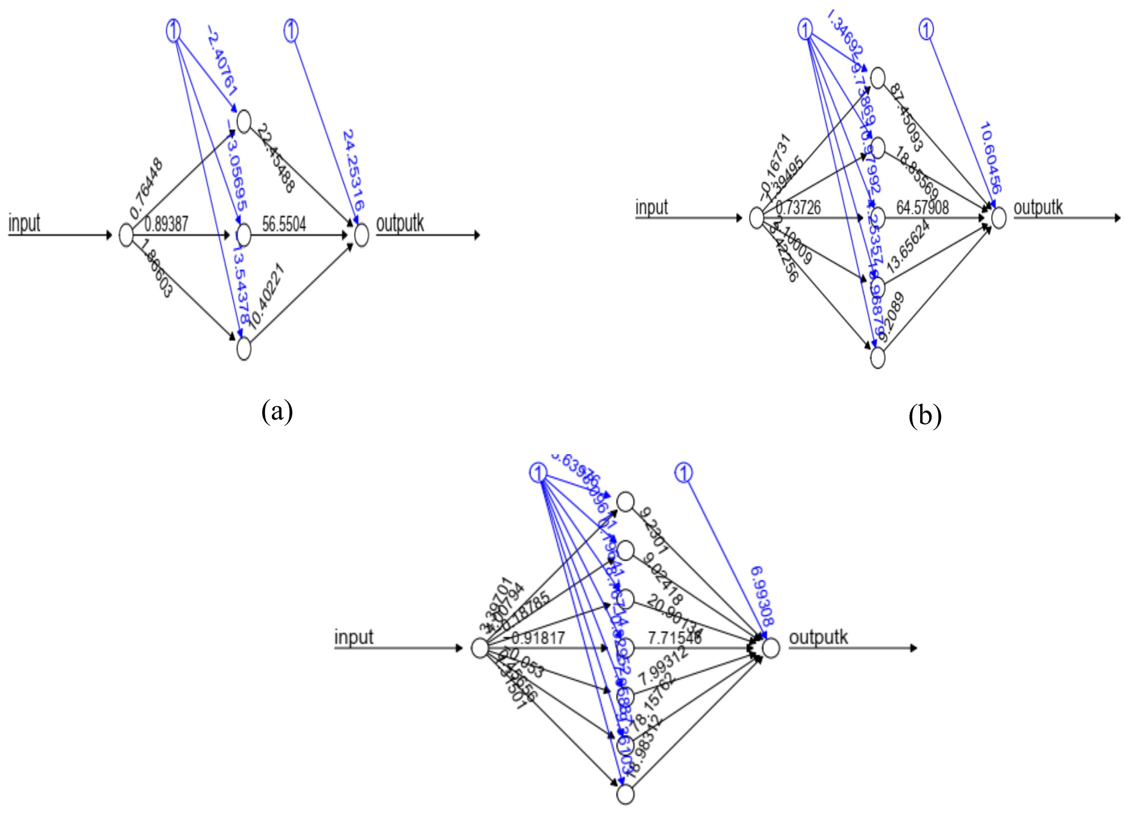
Table 2. RMSE value in each treatment with hidden layer variations

Treatment	RMSE		
	1-3-1	1-5-1	1-7-1
control	0.294	0.140	0.157
50% shade	0.160	0.161	0.165
80% shade	0.321	0.315	0.082

The RMSE results from the three treatments show differences in the values of each architecture. The smallest RMSE value was obtained in the architecture using 5 (five) hidden layers in the control treatment, *i.e.*, 0.140. In the 50% shade treatment, the smallest RMSE value was obtained in the architecture using 3 (three) hidden layers, *i.e.*, 0.160. The last treatment is 80% shading, with the smallest RMSE value obtained on the architecture using 7 (seven) hidden layers, which is 0.082. Figure 9 shows a simple ANN architecture with hidden layer variations.

In the case of simple ANN used in this study, there was no difference between the data being trained and the data being tested because only 12 plant growth data were used. Likewise, the normalization and denormalization processes are also not used in this study because the data used as input has the same degree as the output data. It was explained that normalization is needed if input and output data are not uniform, so it needs to be converted into an index whose values range from 0 to 1 [30].

The comparison using the regression model with the ANN model showed that the ANN model was better at predicting the growth of red chili plants. The average RMSE value in the ANN model of 0.127 is smaller than the average RMSE value in the 2nd-order polynomial regression model of 1.53. The following is a comparison of the actual plant growth value with the predicted plant growth from the results of the ANN model, as presented in Table 3.



(c)

Figure 9 ANN model of various hidden layers: (a) three hidden layers, (b) five hidden layers, (c) seven hidden layers

Table 3. Data of actual and predicted plant growth

Week	Actual plant height (cm)			Predicted plant height (cm)		
	control	50% shade	80% shade	control	50% shade	80% shade
1	27.8	28.3	28.8	27.80	28.30	28.80
2	31.0	31.3	32.4	31.00	31.28	32.39
3	34.9	35.1	35.8	34.89	35.07	35.79
4	38.6	38.9	38.5	38.60	39.09	38.49
5	43.1	42.6	44.0	43.07	42.17	43.90
6	44.9	44.3	46.2	44.97	44.78	46.20
7	49.9	48.1	50.6	49.73	47.78	50.60
8	54.3	50.8	53.8	54.63	50.92	53.67
9	57.4	53.3	55.6	56.85	53.24	55.91
10	57.4	54.5	57.9	57.82	54.49	57.69
11	59.3	54.9	59.1	59.18	55.05	58.94
12	62.1	55.4	59.6	62.10	55.27	59.76

The ANN model's coefficient of determination value (R^2) in control, 50% shade, and 80% shade treatment was 0.99, 0.99, and 0.99, respectively. From the three treatments, the R^2 value is close to 1; this result means that the ANN model is perfect for predicting plant growth. However, to produce a better result, based on previous research, the genetic neural network prediction model could be used to improve the neural network prediction model [31].

4. CONCLUSION

Plant growth prediction can be modeled using regression models and ANN models. The results showed that the ANN model had an average RMSE value of 0.127, and the average RMSE value of the regression model was 1.53. The smaller the RMSE value, the better the model, so the ANN model is better at predicting plant growth than the regression model.

AUTHORS' CONTRIBUTIONS

Guyup Mahardhian Dwi Putra prepared the literature study, modified the program code, analyzed data, and wrote the manuscript. Lilik Sutiarso designed the research model system. Andri Prima Nugroho developed the program code and reviewed the manuscript. Ngadisih created the research environment settings. Salman Ibnu Chaer designed the control system.

ACKNOWLEDGMENTS

This study was supported by the Department of Agricultural and Biosystems Engineering, Faculty of Agricultural Technology at the Universitas Gadjah Mada, Smart Agriculture Research group of Agricultural and Biosystems Engineering UGM for the support, and also Indonesia Endowment Fund for Education (LPDP) Ministry of Finance Republic Indonesia

REFERENCES

[1] A. H. Nasution, Hanter, P. Rahman, Marketing efficiency of red chilli pepper in North Sumatera Province, IOP Conf. Ser. Earth Environ. Sci., 782(2) (2021). DOI: https://doi.org/10.1088/1755-315/782/2/022029.

[2] S. Kusumaningrum, E. Sulistyaningsih, R. Harimurti, K. Dewi, Physiological characters of four lowland chilli varieties (*Capsicum annum* L.) with root cutting, IOP Conf. Ser. Earth Environ. Sci., 883(1) (2021). DOI: https://doi.org/10.1088/1755-315/883/1/012073.

[3] F. B. Salisbury, C. W. Ros, *Plant Physiology*. 1992.

[4] X. Zhang, S. You, Y. Tian, J. Li, Comparison of plastic film, biodegradable paper and bio-based film mulching for summer tomato production: Soil properties, plant growth, fruit yield and fruit quality, Sci. Hortic. (Amsterdam)., 249(489) (2019) 38–48. DOI: https://doi.org/10.1016/j.scienta.2019.01.037.

[5] P. Kalaitzoglou, W. van Ieperen, J. Harbinson, M. van der Meer, S. Martinakos, K. Weerheim, C. C. S. Nicole, L. F. M. Marcelis, Effects of continuous or end-of-day far-red light on tomato plant growth, morphology, light absorption, and fruit production, Front. Plant Sci., 10(March) (2019) 1–11. DOI: https://doi.org/10.3389/fpls.2019.00322.

[6] J. Li, Y. Gao, X. Zhang, P. Tian, J. Li, Y. Tian, Comprehensive comparison of different saline water irrigation strategies for tomato production: Soil properties, plant growth, fruit yield and fruit quality, Agric. Water Manag., vol. 213(489) (2019) 521–533. DOI: https://doi.org/10.1016/j.agwat.2018.11.003.

[7] C. Zisis, E. M. Pechlivani, S. Tsimikli, E. Mekeridis, A. Laskarakis, S. Logothetidis, Organic photovoltaics on greenhouse rooftops: Effects on plant growth, Mater. Today Proc., volume 21 (2020) 65–72. DOI: https://doi.org/10.1016/j.matpr.2019.07.658.

[8] A. Rizkiana, A. P. Nugroho, M. A. Irfan, L. Sutiarso, T. Okayasu, Crop growth prediction model at vegetative phase to support the precision agriculture application in plant factory, AIP Conf. Proc., volume 2202, (2019). DOI: https://doi.org/10.1063/1.5141717.

[9] J. K. Basak, W. Qasim, F. G. Okyere, F. Khan, Y. J. Lee, J. Park, H. T. Kim, Regression Analysis to Estimate Morphology Parameters of Pepper Plant in a Controlled Greenhouse System, J. Biosyst. Eng., 44(2) (2019) 57–68. DOI: https://doi.org/10.1007/s42853-019-00014-0.

[10] A. Rizkiana, A. P. Nugroho, N. M. Salma, S. Afif, R. E. Masithoh, L. Sutiarso, T. Okayasu, Plant growth prediction model for lettuce (Lactuca sativa.) in plant factories using artificial neural network, IOP Conf. Ser. Earth Environ. Sci., volume 733(1) (2021). DOI: https://doi.org/10.1088/1755-315/733/1/012027.

[11] M. Niazian, S. A. Sadat-Noori, M. Abdipour, Artificial neural network and multiple regression analysis models to predict essential oil content of ajowan (Carum copticum L.), J. Appl. Res. Med. Aromat. Plants, volume 9 (2018) 124–131. DOI: https://doi.org/10.1016/j.jarmap.2018.04.001.

[12] Y. Astutik, Murad, G. M. D. Putra, D. A. Setiawati, Remote monitoring systems in greenhouse based on NodeMCU ESP8266 microcontroller and Android, AIP Conf. Proc., volume 2199 (2019) 1–14. DOI: https://doi.org/10.1063/1.5141286.

[13] A. Priyati, M. Sahbandi, G. M. D. Putra, D. A. Setiawati, The Design of Automatic Sprinkler based on Arduino Uno Microcontroller, IOP Conf. Ser. Earth Environ. Sci., 355(1) (2019) DOI: https://doi.org/10.1088/1755-1315/355/1/012088.

[14] Y. H. Hu, S. C. Yu, X. Qi, W. J. Zheng, Q. Q. Wang, H. Y. Yao, An overview of multiple linear regression model and its application, Zhonghua Yu Fang Yi Xue Za Zhi, 53(6) (2019) 653–656.

[15] U. Simonsohn, Two Lines: A Valid Alternative to the Invalid Testing of U-Shaped Relationships With Quadratic Regressions, Adv. Methods Pract. Psychol. Sci., 1(4) (2018) 538–555. DOI: https://doi.org/10.1177/2515245918805755.

[16] J. M. Cerqueira, J. M. Moura, C. Sylla, L. Ferreira, An Augmented Reality Mathematics Serious Game, OpenAccess Ser. Informatics, 81(6) (2020) 1–6. DOI: https://doi.org/10.4230/OASIcs.ICPEC.2020.6.

[17] E. A. Graf, J. H. Fife, H. Howell, E. Marquez, The Development of a Quadratic Functions Learning Progression and Associated Task Shells, ETS Res. Rep. Ser., 2018(1) (2018) 1–28. DOI: https://doi.org/10.1002/ets2.12234.

[18] G. M. D. Putra, D. A. Setiawati, Murad, Study of mathematics modeling on ginger geometric changes during drying using image analysis, IOP Conf. Ser. Earth Environ. Sci., 733(1), (2021). DOI: https://doi.org/10.1088/1755-1315/733/1/012002.

[19] N. A. Al-Sammarraie, Y. M. H. Al-Mayali, Y. A. Baker El-Ebiary, Classification and diagnosis using back propagation Artificial Neural Networks (ANN) algorithm, 2018 Int. Conf. Smart Comput. Electron. Enterp. ICSCEE 2018, 2018, pp. 1–5, DOI: https://doi.org/10.1109/ICSCEE.2018.8538383.

[20] Z. Wan, J. Li, L. Shu, M. Luo, X. Li, S. Fu, K. Xu, Nonlinear equalization based on pruned artificial neural networks for 112-Gb/s SSB-PAM4 transmission over 80-km SSMF, Opt. Express, 26(8) (2018) 10631. DOI: https://doi.org/10.1364/oe.26.010631.

[21] H. Jahangir, H. Tayarani, A. Ahmadian, M. A. Golkar, J. Miret, M. Tayarani, H. O. Gao, Charging demand of Plug-in Electric Vehicles: Forecasting travel behavior based on a novel Rough Artificial Neural Network approach, J. Clean. Prod., 229(2019) (2019) 1029–1044. DOI: https://doi.org/10.1016/j.jclepro.2019.04.345.

[22] V. Wagh, D. Panaskar, A. Muley, S. Mukate, S. Gaikwad, Neural network modelling for nitrate concentration in groundwater of Kadava River basin, Nashik, Maharashtra, India, Groundw. Sustain. Dev., volume 7 (2018) 436–445. DOI: https://doi.org/10.1016/j.gsd.2017.12.012.

[23] R. Kumar, P. Kumar, S. Singh, R. Vijay, Fast and accurate synthesis of frequency reconfigurable slot antenna using back propagation network, AEU - Int. J. Electron. Commun., volume 112 (2019) 1–10. DOI:

https://doi.org/10.1016/j.aeue.2019.152962.

[24] W. Wang, Y. Lu, Analysis of the Mean Absolute Error (MAE) and the Root Mean Square Error (RMSE) in Assessing Rounding Model, IOP Conf. Ser. Mater. Sci. Eng., 324(1) (2018). DOI: https://doi.org/10.1088/1757-899X/324/1/012049.

[25] V. V. Kusnetsov, A. S. Doroshenko, N. V. Kudryakova, M. N. Danilova, Role of Phytohormones and Light in De-etiolation, Russ. J. Plant Physiol., 67(6) (2020) 971–984. DOI: https://doi.org/10.1134/S1021443720060102.

[26] T. Armarego-Marriott, O. Sandoval-Ibañez, Ł. Kowalewska, Beyond the darkness: Recent lessons from etiolation and de-etiolation studies, J. Exp. Bot., 71(4) (2020) 1215–1225. DOI: https://doi.org/10.1093/jxb/erz496.

[27] D. Floris, W. Kühlbrandt, Molecular landscape of etioplast inner membranes in higher plants, Nat. Plants, 7(4) (2021) 514–523. DOI: https://doi.org/10.1038/s41477-021-00896-z.

[28] I. Muda, A. Haris Harahap, E. Erlina, S. Ginting, A. Maksum, E. Abubakar, Factors of quality of financial report of local government in Indonesia, IOP Conf. Ser. Earth Environ. Sci., 126(1) (2018). DOI: https://doi.org/10.1088/1755-1315/126/1/012067.

[29] O. G. Ogunyale, O. O. Fawibe, A. A. Ajiboye, D. A. Agboola, A Review of Plant Growth Substances: Their Forms, Structures, Synthesis and Functions, 5(4) (2014), 152–168.

[30] J. H. Jung, H. Chung, Y. S. Kwon, I. M. Lee, An ANN to Predict Ground Condition ahead of Tunnel Face using TBM Operational Data, KSCE J. Civ. Eng., 23(7) (2019) 3200–3206. DOI: https://doi.org/10.1007/s12205-019-1460-9.

[31] J. Hu, P. Xin, S. Zhang, H. Zhang, D. He, Model for tomato photosynthetic rate based on neural network with genetic algorithm, Int. J. Agric. Biol. Eng., 12(1) (2019) 179–185. DOI: https://doi.org/10.25165/j.ijabe.20191201.3127.