❐   3562

Growth predictions of lettuce in hydroponic farm using autoregressive integrated moving average model

Muhammad Zacky Asy'ari[1], Julius Felix Chrysdi Aten[1], Dimas Prasetyo[2]

[1]Automotive and Robotic Engineering, Department of Computer Science, Binus ASO School of Engineering, Binus University, Jakarta, Indonesia

[2]Department of Industrial Engineering, Faculty of Engineering, Binus University, Jakarta, Indonesia

Article Info

ABSTRACT

Hydroponic farming techniques can grow plants faster when intensive monitoring is carried out. However, sometimes proper monitoring does not occur, which causes some plants not to grow as expected. This research uses a sensor to measure a parameter that affects the growth of the hydroponic plants and uses a camera to take pictures periodically to measure daily growth. The research presented in this article is to build a model and forecasting plant growth in hydroponics farming using a time series approach. This paper demonstrated that the historical data on lettuce growth could be used to predict future plant size. The autoregressive integrated moving average (ARIMA) model was used and analyzed according to the six performance criteria: mean absolute error (MAE), root mean squared error (RMSE), mean absolute percentage error (MAPE), Akaike information criterion (AIC), and Bayesian information criterion (BIC). To find the best model, different autoregressive (p) and moving average (q) parameters were examined. We find that the appropriate statistical model for lettuce growth is ARIMA (2, 2, 1) which has the lowest AIC, BIC, and MAPE with values of 76.67, 79.02, and 0.04, respectively, to forecast the plant size for the next three days.

Corresponding Author:

Muhammad Zacky Asy'ari
Automotive and Robotic Engineering, Department of Computer Science
Binus ASO School of Engineering, Binus University
Tangerang Selatan, Banten, Jakarta, Indonesia
Email: muhammad.zacky.asyari@gmail.com

1. INTRODUCTION

Agriculture is the second largest supporting sector for Indonesia's economy, which is strongly related to the food demand in the country. The purchasing power of the people in Indonesia greatly influences food security. The impact of COVID-19 has dramatically affected the economic situation in the country, so controlling food prices is one of the strategies to overcome the problem [1]. There are several alternatives for the household to support their daily needs by harvesting foods on their own, particularly vegetables. Hydroponics is one of the farming techniques that is often used by modern society. Because of their limited land, this farming method has several advantages, such as not requiring a large area, growing faster, and having higher quality yields. Nevertheless, several parameters play a vital role in determining harvesting success when farming using hydroponic techniques, such as total dissolved solids (TDS), pH, humidity and air temperature, solution temperature, and light intensity [2]. If one or more of these parameters are not suitable for the growth of hydroponic lettuce plants, it will cause them to grow abnormally and even die [3].

IoT solutions in many industrial environments can lead to innovative, productive, and precise systems development to increase efficiency in every intelligent operation. The IoT and computer vision approaches combined with machine learning (ML) algorithms and data mining to evaluate the system's performance are used by researchers. Computer vision has different applications in smart farming, especially in hydroponics, such as monitoring plant growth, estimating crop yield, and measuring nutrient constituents. Develop a model in computer vision using ML algorithms to instruct computers to perform complex tasks through regression, diagnosis, planning, and recognition by learning from historical data. Thus, data and algorithms are considered fundamental to the performance of ML models. High quality data and more extensive data sizes are essential for ML model accuracy.

For example, many researchers utilize computer vision technology in agriculture to recognize crops that benefit from non-destructive and convenient characteristics [4]. This study has undiscovered techniques to improve, especially real time detection speed and dead seedling observation. The most popular features of the computer vision technique are image recognition, texture, color extraction, shape identification, and other characteristics of the picture. In addition, it distinguishes targets in terms of different characteristics.

This paper proposes an IoT hydroponic plant monitoring system that utilizes a few sensors and a camera as input data to measure and predict plant growth parameters. Using a time series forecasting algorithm called autoregressive integrated moving average (ARIMA), the data will also be used to forecast hydroponic plant growth for the next few days. The performance of each algorithm will be evaluated with the mean absolute error (MAE), root mean squared error (RMSE), and mean absolute percentage error (MAPE) value [5]. On the other hand, the ARIMA model complexity measure performance is evaluated with Akaike information criterion (AIC) and Bayesian information criterion (BIC) [6]. In the end, the model with the best performance will be used to predict plant growth which aims to provide information to hydroponic farmers so that they can take preventive measures to produce higher quality crops.

2. LITERATURE REVIEW

Over the past few years, there has been a development in ML assisted agriculture as part of applications in the IoT for precision agriculture. This interest is driven by the benefits of the practice in increasing agricultural productivity, sustainability, and profitability while increasing food security [7]. As a result, the ML algorithm has gained attention in many fields. The most known algorithms include linear models, support vector machines, clustering, decision trees, random forests, neural networks, and clustering. Additionally, in computer vision, data sources such as RGB color, visible light, thermal infrared image, near 3D, and spectroscopy to measure and analyze features like texture, shape, spectrum, and color.

A combination method on the regional center of cross-border leaves and a methodology to improve watershed segmentation on overlapping leaf images were developed to find the leaf area on every seedling in the plug tray [8]. This vision system was developed to measure leaf area in each leaf cell to distinguish good leaves based on proportion in area. The top view of the seedlings and the method for calculating each seedling leaf area in the plug tray were investigated and the detection process was developed. On the other hand, a crop segmentation method (AP-HI) was used to automatically detect two critical growths of maize seedlings. The result gives a high performance of 96.68% and could outstand exterior environment conditions [9]. Meanwhile, over-segmentation and sensitivity to false edges were the limitations of this method.

Furthermore, according to the feature position, the over segmentation block projection technique was utilized to locate the crown of maize seedlings captured by the camera. The camera image plane is parallel to the ground. Hence, the center of maize seedling roots was obtained. Arbitrary, although these approaches give high efficiency and uncomplicated calculation, they are unsuitable for diversified environmental conditions and destitute versatility. Moreover, they have low robustness and are sensitive to noise with low robustness [10].

The combination of time series-forecasting algorithms and computer vision has become one of the solutions to overcoming problems in the agricultural sector, especially in forecasting growth and the variables that affect growth. For example, research conducted by Srivani et al. [11] used the long short term memory (LSTM) algorithm to predict the value of the root zone temperature (RZT) in an indoor hydroponic. The main contribution of this study is to analyze which hyperparameter combination shows better performance with the smallest average RMSE error. The study aims to design a model that can predict environmental changes so that it can adapt and control the actuators automatically. The use of computer vision can also be used to see the response of plants under a given environmental condition. According to Story and Kacira [12], utilize cameras and computer vision to continuously monitor color, morphological, textural, and spectral (crop indices and temperature) features from a crop that can monitor plant growth and health status and improve controlled environmental conditions results.

Our research paper is conducted on hydroponic lettuce seedlings as the objective and proposes forecasting hydroponic plant growth for the next few days. A time series forecasting algorithm called ARIMA was used to detect them. To construct the dataset, the prediction model can better recognize the growth of hydroponic lettuce through computer vision data extraction.

3.   METHOD

This study uses two main concepts for data collection: machine to machine (M2M) communication and the IoT. M2M allows a machine to communicate with other machines without requiring human interference. IoT also supports this concept where the machine must be connected to the internet network to communicate or exchange data [13]. These two concepts are used explicitly on a Raspberry Pi 3B+ and a Wemos D1 R32 (microcontroller) connected to various sensors in collecting all the data needed in this study.

3.1. Device setup

The Raspberry Pi 3B+ is used as a web server where the Thingsboard IoT platform is installed with the addition of a PostgreSQL database to store all the sensor data [14]. On the other hand, a small yet powerful Raspberry Pi camera module is also added. The camera automatically takes photos of plant growth every four hours. The sensors used in this experiment are pH, TDS, temperature, and humidity, which are directly connected to the microcontroller as illustrated on Figure 1(a).

The device is connected to the local internet and sends data every 2 minutes into the PostgreSQL database. An access point on the system also to communicates with the Raspberry Pi and the microcontroller Wemos D1 R32 is used for data acquisition, as presented in Figure 1(b). Hence, they can exchange data wirelessly within the intranet network.
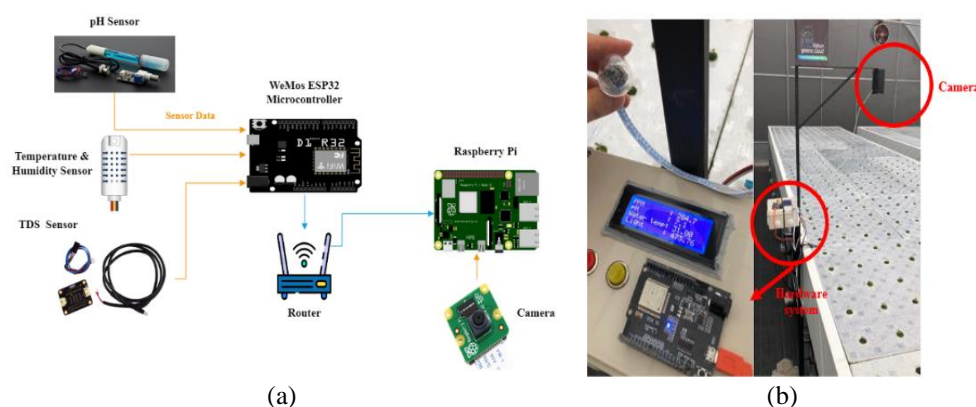

(a)   (b)

Figure 1. IoT infrastructure design breakdown into; (a) system architecture and (b) hardware setup

3.2. Data collection

Data were collected at the uncontrolled greenhouse for eight days from 8–15 June 2022. Sample images are taken every 3 hours. However, the data at night does not appear to be used in the analysis because greenhouse lighting depends only on sunlight. There are no additional lights for lighting during nighttime. While the data from the sensors to measure pH, TDS, temperature, and humidity were recorded every 5 minutes.

3.3. Image collection and image processing

Photo collection, as seen in Figure 1(b) using an additional camera module connected to a Raspberry Pi, is intended to show plant growth every hour. The photo will also undergo several image processing to extract the leaf area from the photo. In this study, a software called ImageJ. ImageJ (fiji) is a Java based image processing program developed by the National Institutes of Health and the Laboratory for Optical and Computational Instrumentation. ImageJ can read various image file formats, including PNG, GIF, JPEG, DICOM, FITS, and RAW. Some of the uses of this program are to display, edit, calibrate, measure, and analyze image data. ImageJ is often used in various fields, such as biology, earth science, fluid dynamics, astronomy, computer vision, and signal processing [15].

The ImageJ software is intended to measure the area (area) of lettuce leaves based on photos captured by the Raspberry Pi. The algorithm used by ImageJ in measuring is a light threshold and pixel counting. The light threshold is done by converting the image into an 8-bit format (grayscale) and maximizing the image's contrast to distinguish between lettuce leaf objects and the image's background. The pixel counting calibration uses an object whose size is known by manual measurement (such as using a ruler). In the next step, ImageJ calculates pixels number per centimeter as a reference to measure the lettuce leaf area.

The ImageJ software can perform image processing series such as editing, calibrating, measuring, and analyzing image data. Two methods to convert image data into leaf area are pixel counting and light threshold. Pixel counting is a method to measure an object according to another object's length. As a result, ImageJ can easily find the scale that will be the reference in pixels per centimeter. The light threshold is the other method to adjust an image's color in RGB format. This process aims to separate the leaf object from its background. Therefore, the region of interest (ROI) selection is accurate. Figure 2 is an example of output from the ImageJ software.
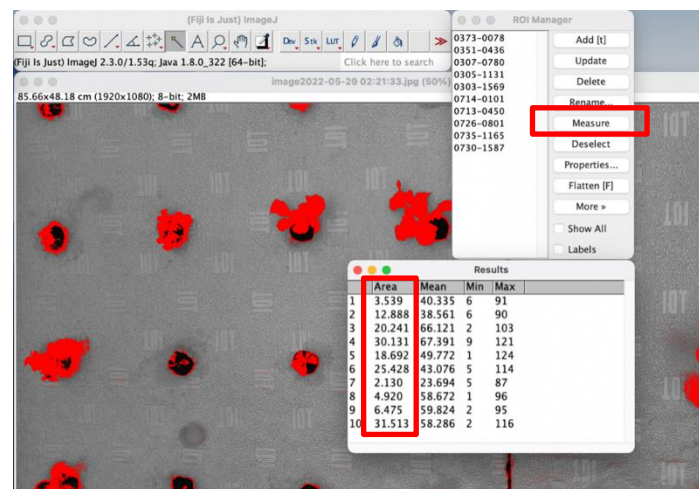

Figure 2. The light threshold to separate the lettuce leaf object from the background

According to the camera's photo, ten specimens or samples are present in the frame. Sample number 3 in Figure 3 has a significant size compared to the others. Hence, it is interesting to monitor its plant growth to be used as a reference point in the dataset.
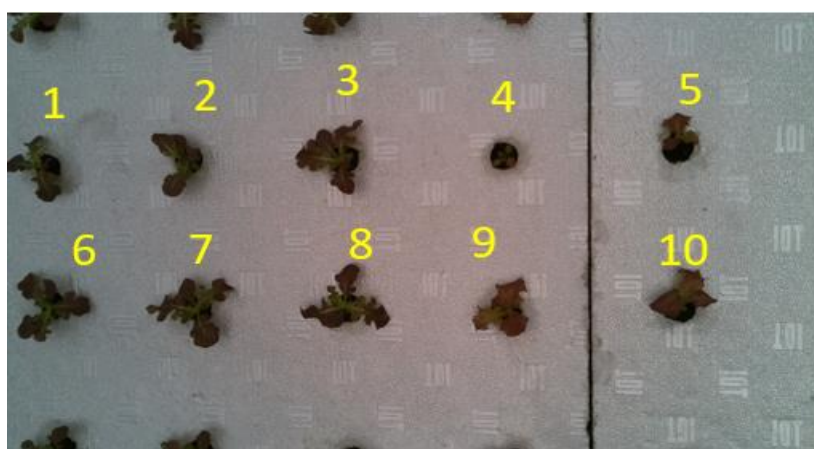

Figure 3. Sampling on the dataset

The line chart of lettuce leaf area growth in the dataset shows that sample 3 had the large leaf area on day eight among the best four samples, as illustrated in Figure 4. In contrast, the other samples did not experience significant growth. Therefore, data from sample 3 will be used for time series forecasting models in the following part of this paper. The uneven distribution of nutrient fluids can cause a difference in growth, so hydroponic plant seeds' absorption is not optimal for each sample [16].
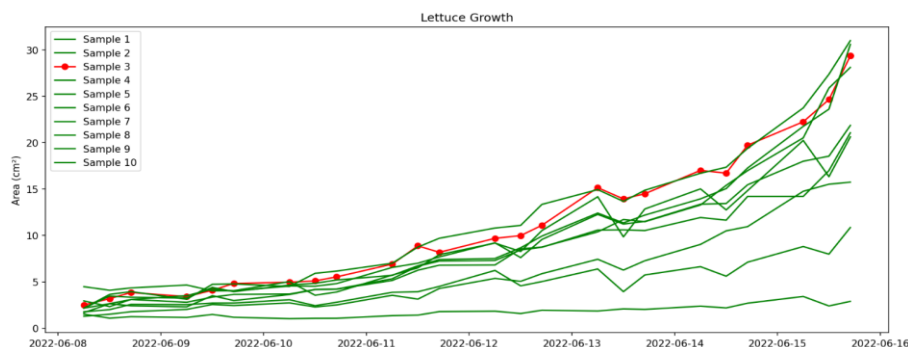

Figure 4. Time series hydroponics lettuce growth area

3.4. Time series forecasting model

The model creation process is carried out using one of the Python statistics packages. The process begins with testing the dataset's stationarity and determining the model's order and lag using auto correlation function (ACF) and partial auto correlation function (PACF) plots [17]. First, the model was built using the time series data (sensor and image). Then, the dataset stationarity test was carried out using an augmented dickey fuller (ADF) statistical test [18].

The parameter that becomes the reference for determining the stationarity of a dataset is the p-value. If the p-value is less than 0.05, then the dataset is stationary. The dataset is not stationary if the p-value is more significant than 0.05 [19]. When the dataset is not stationary, it can be differentiated by subtracting each value from the previous value (t–(t-1)) [20]. The goal is to make the mean, variance, and standard deviation constant.

ARIMA model is a time series forecasting algorithm that uses one variable (univariate) in modeling without considering other variables. ARIMA combines auto regressive (AR) and moving average (MA) [21]. Both models use the historical value to make a prediction and previous error values to make a prediction, respectively. The autoregressive integrated moving average with exogenous variables (ARIMAX) is the ARIMA model with additional multivariate properties [22]. The ARIMAX model considers external or exogenous variables to predict future data.

The seasonal auto-regressive integrated moving average (SARIMA) model is a time series forecasting algorithm aimed at forecasting datasets that have seasonal patterns [23]. A seasonal pattern can be defined as a pattern that keeps repeating itself over a particular time, for example, daily, weekly, or even yearly. The ARIMA model's parameters are (p, d, q). In contrast, in the SARIMA model, the parameters used are (p, d, q) (P, D, Q) m. The additional parameters in the form of (P, D, Q) m are used to determine AR, differentiation, and MA orders from the SARIMA seasonal pattern. The SARIMAX is the SARIMA model with more than one non-independent variable (multivariate). The SARIMAX model considers external or exogenous variables to predict future data [24].

3.5. Auto correlation function and partial auto correlation function

ACF and PACF plots are used to determine orders from ARIMA, ARIMAX, SARIMA, and SARIMAX models for both seasonal and non-seasonal orders. The ACF plot is used to determine the order of the MA, which measures the direct and indirect effects on the dataset at a certain lag [25]. At the same time, the PACF plot is used to determine the order of AR, which only measures the direct effect on the dataset at a certain lag.

3.6. Determining the model

The first step in making a model for time series forecasting is a stationarity test of the dataset. It used a statistical test called ADF. It is an indicator of whether a dataset is stationary through the p-value. The dataset became stationer if the p-value was below 0.05. On the contrary, if above its minimum value, the

dataset should be differentiated to eliminate the existing trend and value variance. It occurs by subtracting each value from the previous value (t–(t-1)).

The dataset was analyzed to illustrate the process. After the differentiation process, the p-value decreased until equal to 0.0013. Due to the differentiation process being twice, the dataset's order d equals 2. The following step is to determine the p and q values. It is observed that the PACF lag 1 is quite significant since it is well above the significance line; hence the p AR value is 1. On the other hand, that 1 of the lags is out of the significance limit in the ACF plot, so it is said that the optimal value of q MA is 1.

In this paper, we experimented by varying the values of p and q from 1 to 3. In contrast, the value of d for differentiating is 2. Next, we looked at the performance of each model with these different parameters and combinations. In addition, specify the model's order using ACF plots and PACF plots. A python library called 'pmd_arima' can be used in a function [26], [27].

The function inside the library package called 'auto_arima' can automatically generate models with various order combinations (p, d, q) [28]. This function will provide output and information in the form of a model with the smallest error value and its order. The results of the auto_arima function obtained the order (p, d, q) is equal to 2, 2, 1. The auto_arima function provides information that the model with the best order combination produces the minor error value.

3.7. Evaluation criteria

Several criteria need to be evaluated to determine the model with the slightest error. The criteria that are used for ARIMA models in this study are RMSE, MAE, MAPE, AIC, Akaike information criterion bias corrected (AICC), and BIC. The formulas to calculate each parameter are stated in Table 1.

Table 1. Performance criteria calculation

Performance criteria	
RMSE	AIC
$RMSE = \sqrt{\dfrac{1}{n}\sum_{t=1}^{n}|Y_t - \hat{Y}_t|^2}$	$AIC = -2\ln(\theta) + 2k$
MAE	AICC
$MAE = \left\|\dfrac{(Y_t - \hat{Y}_t)}{n}\right\|$	$AICc = AIC + \dfrac{2(k+1) + 2(k+2)}{n - k - 2}$
MAPE	BIC
$MAPE = \dfrac{1}{n}\sum_{t=1}^{n}\left\|\dfrac{Y_t - \hat{Y}_t}{Y_t} \times 100\right\|$	$BIC = -2\ln(\theta) + k\ln(n)$

Where Θ is maximum likelihood function, n is observation number, k is a number of model parameters, Y_t is observation time, and $\hat{Y}_t$ is estimated observations values.

4. RESULTS AND DISCUSSION

The statistical results of the evaluation criteria from several models are compared, such as the mean, standard deviation, minimum, and minimum values presented in Table 2. Based on the analysis, results from the standard deviation for all parameters are relatively low, meaning that the data are clustered around the mean. Therefore, the selected ARIMA model comparison has minor significant differences.

Table 2. Statistical summary on accuracy selection criteria

	RMSE	MAE	MAPE	AIC	AICC	BIC
Mean	1.326720	1.258970	0.053110	80.88573	84.940095	86.340950
Std	0.361076	0.306662	0.012831	2.998304	3.196911	2.874991
Min	0.974652	0.946650	0.040245	76.672392	79.025333	81.036561
Max	1.867261	1.728319	0.071706	85.724950	89.267932	90.089120

According to Figure 5(a), it can be seen that the most significant error of all performance comparison parameters is ARIMA (1, 2, 1), where this parameter is the first guess based on looking at ACF and PACF in the previous section. Because it is necessary to iterate over the ARIMA parameters with different parameters, it is found that the results that have a minor error are ARIMA (2, 2, 1) and auto_arima also recommends these parameters. In Figure 5(b), ARIMA (2, 2, 1) shows consistency by having the lowest value on the compared three parameters. Hence, it is the most suitable parameter for the lettuce hydroponics plant dataset. Later, the model will forecast the development of lettuce plants on a hydroponic system.
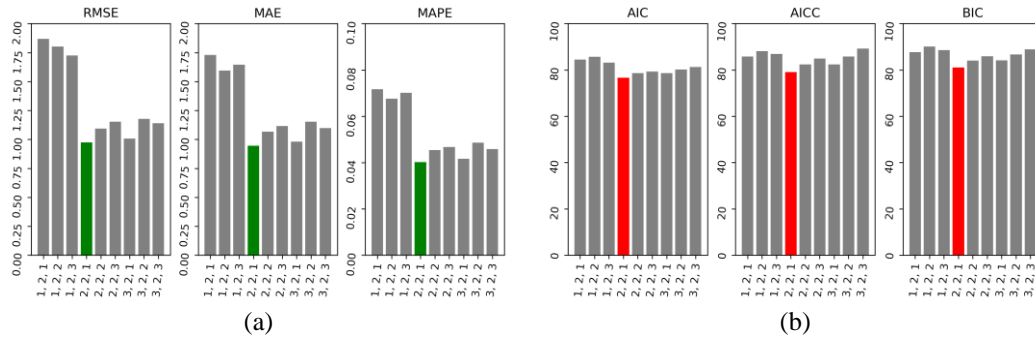
Figure 5. Simulation of result; (a) accuracy measure and (b) complexity of the ARIMA model

Since ARIMA (2, 2, 1) model fits the lettuce growth data, as shown in Figure 6(a), the trend between measured data and predictions is matched. Therefore, it can directly forecast the area for the next three days out of eight days (three samples each day). Figure 6(b) presents the actual and forecasted area in cm^2 with a 95% confidence limit.

The forecasted values indicate that the lettuce hydroponics growth will continue to rise. Keep in mind that the result is a predicted value, but the hydroponic planting is a dynamic that depends on the surrounding environment. Hence, we should pay attention to climatic conditions, nutrition liquids, and pest attacks. Proper adjustment and decision-making in operation are necessary to maintain the growth trend and control to prevent crop failure.
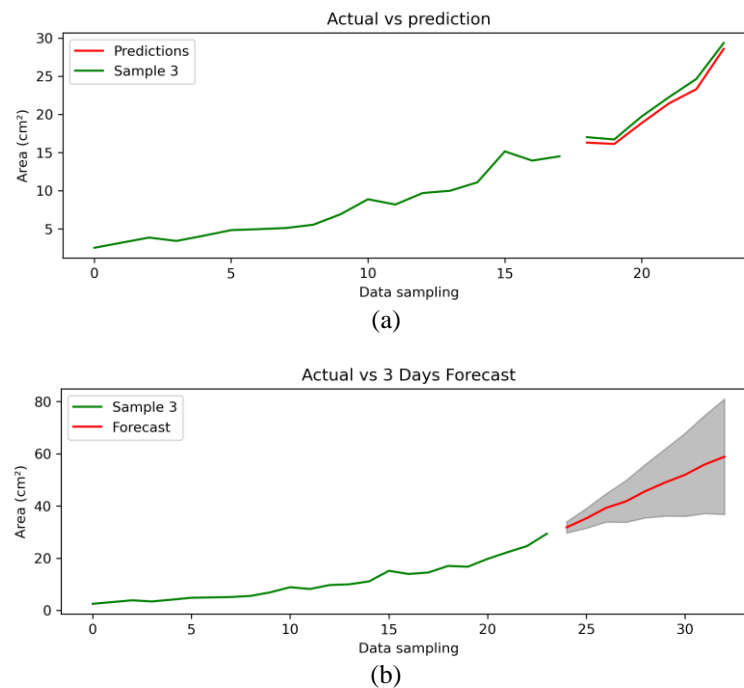

(a)

(b)

Figure 6. The result of ARIMA model in; (a) actual vs prediction plot and (b) three days ahead forecast plot

5. CONCLUSION

This paper discusses finding the best ARIMA parameter and forecasting hydroponic lettuce growth based on eight days of data. Time series plots, ACF, and PACF, were used to test data stationarity. The ARIMA model with different orders of AR (p) and MA (q) were compared. The best model is ARIMA (2, 2, 1) because it has a minimum value for all compared performance criteria.

The performance ARIMA (2, 2, 1) time series forecasting model in predicting hydroponic plants' growth gives the smallest value of RMSE, MAE, and MAPE with 0.97, 0.94, and 0.04, respectively. The real time monitoring using sensors and cameras will significantly ease the work of farmers. Another benefit is that the system can predict plant growth over the next few days. Therefore, it can be used as a reference to make an early decision to produce high quality crops.

REFERENCES

[1] S. Arif, W. Isdijoso, A. R. Fatah, and A. R. Tamyis, "Strategic Review of Food Security and Nutrition in Indonesia," *The SMERU Research Institute: Jakarta*, pp. 38–41, 2020.

[2] M. P. N. Gent, "Factors Affecting Relative Growth Rate of Lettuce and Spinach in Hydroponics in a Greenhouse," *HortScience*, vol. 52, no. 12, pp. 1742–1747, Dec. 2017, doi: 10.21273/HORTSCI12477-17.

[3] C. R. Adams, M. P. Early, and K. M. Bamford, *Principles of Horticulture*. Routledge, 2008, doi: 10.4324/9780080942964.

[4] Z. Li *et al.*, "A high-precision detection method of hydroponic lettuce seedlings status based on improved Faster RCNN," *Computers and Electronics in Agriculture*, vol. 182, pp. 1–11, Mar. 2021, doi: 10.1016/j.compag.2021.106054.

[5] W. Wang and Y. Lu, "Analysis of the Mean Absolute Error (MAE) and the Root Mean Square Error (RMSE) in Assessing Rounding Model," *IOP Conference Series: Materials Science and Engineering*, vol. 324, no. 1, pp. 1–11, Mar. 2018, doi: 10.1088/1757-899X/324/1/012049.

[6] M. J. Brewer, A. Butler, and S. L. Cooksley, "The relative performance of AIC, AICC and BIC in the presence of unobserved heterogeneity," *Methods in Ecology and Evolution*, vol. 7, no. 6, pp. 679–692, 2016, doi: 10.1111/2041-210X.12541.

[7] R. Gebbers and V. I. Adamchuk, "Precision Agriculture and Food Security," *Science*, vol. 327, no. 5967, pp. 828–831, Feb. 2010, doi: 10.1126/science.1183899.

[8] J. H. Tong, J. B. Li, and H. Y. Jiang, "Machine vision techniques for the evaluation of seedling quality based on leaf area," *Biosystems Engineering*, vol. 115, no. 3, pp. 369–379, Jul. 2013, doi: 10.1016/j.biosystemseng.2013.02.006.

[9] Z. Yu *et al.*, "Automatic image-based detection technology for two critical growth stages of maize: Emergence and three-leaf stage," *Agricultural and Forest Meteorology*, vol. 174–175, pp. 65–84, Jun. 2013, doi: 10.1016/j.agrformet.2013.02.011.

[10] Z. Li, R. Guo, M. Li, Y. Chen, and G. Li, "A review of computer vision technologies for plant phenotyping," *Computers and Electronics in Agriculture*, vol. 176, pp. 1–21, Sep. 2020, doi: 10.1016/j.compag.2020.105672.

[11] P. Srivani, C. R. Y. Devi, S. H. Manjula, and K. R. Venugopal, "Time-Series Forecasting and Hyperparameter Tuning with LSTM Approach for Root Zone Temperature of Nutrient Solution In Indoor DWC Hydroponic," *Advance Research in Dynamical & Control Systems*, pp. 962–960, 2020, doi: 10.5373/JARDCS/V12SP6/SP20201114.

[12] D. Story and M. Kacira, "Design and implementation of a computer vision-guided greenhouse crop diagnostics system," *Machine Vision and Applications*, vol. 26, no. 4, pp. 495–506, May 2015, doi: 10.1007/s00138-015-0670-5.

[13] M. Castro, A. J. Jara, and A. F. Skarmeta, "An Analysis of M2M Platforms: Challenges and Opportunities for the Internet of Things," in *2012 Sixth International Conference on Innovative Mobile and Internet Services in Ubiquitous Computing*, IEEE, Jul. 2012, pp. 757–762, doi: 10.1109/IMIS.2012.184.

[14] Thingsboard, "Thingsboard IoT Platform," *Thingsboard.io*. [Online] Available: https://thingsboard.io/. (accessed March. 18, 2022).

[15] T. Ferreira and W. Rasband, "ImageJ User Guide," *IJ 1.46r*, vol. 1, pp. 155–161, 2011.

[16] B. A. Kratky, "Growing lettuce in non-aerated, non-circulated hydroponic systems," *Journal of Vegetable Science*, vol. 11, no. 2, pp. 35–42, 2005, doi: 10.1300/J484v11n02_04.

[17] D. Benvenuto, M. Giovanetti, L. Vassallo, S. Angeletti, and M. Ciccozzi, "Application of the ARIMA model on the COVID-2019 epidemic dataset," *Data in Brief*, vol. 29, pp. 1–4, Apr. 2020, doi: 10.1016/j.dib.2020.105340.

[18] R. Mushtaq, "Augmented Dickey Fuller Test," *SSRN Electronic Journal*, pp. 1–19, 2011, doi: 10.2139/ssrn.1911068.

[19] G. Di Leo and F. Sardanelli, "Statistical significance: p value, 0.05 threshold, and applications to radiomics—reasons for a conservative approach," *European Radiology Experimental*, vol. 4, no. 1, pp. 1–8, Dec. 2020, doi: 10.1186/s41747-020-0145-y.

[20] Q. Wang, S. Li, R. Li, and M. Ma, "Forecasting U.S. shale gas monthly production using a hybrid ARIMA and metabolic nonlinear grey model," *Energy*, vol. 160, pp. 378–387, Oct. 2018, doi: 10.1016/j.energy.2018.07.047.

[21] G. E. P. Box, G. M. Jenkins, G. C. Reinsel, and G. M. Ljung, *Time series analysis: forecasting and control*. Hoboken: John Wiley & Sons, 2015.

[22] W. Anggraeni, K. B. Andri, Sumaryanto, and F. Mahananto, "The Performance of ARIMAX Model and Vector Autoregressive (VAR) Model in Forecasting Strategic Commodity Price in Indonesia," *Procedia Computer Science*, vol. 124, pp. 189–196, 2017, doi: 10.1016/j.procs.2017.12.146.

[23] T. Fang and R. Lahdelma, "Evaluation of a multiple linear regression model and SARIMA model in forecasting heat demand for district heating system," *Applied Energy*, vol. 179, pp. 544–552, Oct. 2016, doi: 10.1016/j.apenergy.2016.06.133.

[24] N. Elamin and M. Fukushige, "Modeling and forecasting hourly electricity demand by SARIMAX with interactions," *Energy*, vol. 165, pp. 257–268, Dec. 2018, doi: 10.1016/j.energy.2018.09.157.

[25] G. Jain and B. Mallick, "A Study of Time Series Models ARIMA and ETS," *SSRN Electronic Journal*, pp. 1–7, 2017, doi: 10.2139/ssrn.2898968.

[26] A. Smith and T. Smith, "Pdmarima Package," *Phyton Package Index*. [Online] Available:

https://pypi.org/project/pmdarima/.(accessed April. 25, 2022).

[27]   T. C. Nokeri, "Forecasting Using ARIMA, SARIMA, and the Additive Model," in *Implementing Machine Learning for Finance*, Berkeley, CA: Apress, 2021, pp. 21–50, doi: 10.1007/978-1-4842-7110-0_2.

[28]   H. Alabdulrazzaq, M. N. Alenezi, Y. Rawajfih, B. A. Alghannam, A. A. Al-Hassan, and F. S. Al-Anzi, "On the accuracy of ARIMA based prediction of COVID-19 spread," *Results in Physics*, vol. 27, pp. 1–17, Aug. 2021, doi: 10.1016/j.rinp.2021.104509.

BIOGRAPHIES OF AUTHORS

Muhammad Zacky Asy'ari received an engineer degree in Mechanical Engineering from Institut Teknologi Bandung (ITB), Indonesia. He received a double master degree in Energy Environment from Ecole Polytechnique, France, and Universitat Politecnica de Catalunya, Spain. Currently, he is a lecturer at Binus University. His research interests include data science, the internet of things, computer vision, and autonomous system. He can be contacted at email: muhammad.zacky.asyari@gmail.com and muhammad.zacky@binus.ac.id.

Julius Felix Chrysdi Aten is an undergraduate student in Automotive and Robotics Engineering from Binus ASO School of Engineering, Binus University, Indonesia. He can be contacted at email: juliusf19@gmail.com.

Dimas Prasetyo is an engineer who obtained his undergraduate study at Industrial Engineering Program at Institut Teknologi Sepuluh Nopember (ITS), Indonesia. Then, he pursued his postgraduate study in University of Indonesia for Master of Engineering and NTU, Taiwan. Currently, he is a lecturer at Binus University. His research interests include optimization and simulation. He can be contacted at email: dimas.prasetyo1@binus.edu.