

Forest Plantation Pest and Disease Forecast Model

Andri Pranolo

Informatics Department
Universitas Ahmad Dahlan
Indonesia
andri.pranolo@tif.uad.ac.id

Siti Muslimah Widyastuti

Faculty of Forestry
Universitas Gadjah Mada
Indonesia
smwidyastuti@ugm.ac.id

Azhari

Computer Science Department
Universitas Gadjah Mada
Indonesia
azhari@ugm.ac.id

Abstract—This research aims to propose a forecast model of pest and disease plantation. The data sample collected by Laboratory of forest health and protection - Faculty of Forestry Universitas Gadjah Mada in the periods of times, so we have time series data of Powdery mildew disease growth which has observed from *Acacia mangium* nursery. This model combined with the expert system model, and the identification and calculation of damage size and the intensity of damage model. The expert system used for identification of pest and disease on plantation, model of identification used for collecting the data and calculating of a size and the intensity of plantation damage, and forecasting model will be used for seeing a disease growing without treatment. In this manuscript, the forecast model is using Auto-Regressive (AR), Moving Average (MA), and ARIMA method to see how the data sample are suitable and the model is working. We are using order (2,1,0) for AR, MA (0,1,2), and combined method ARIMA (2,1,1) which has better RSS value (0.6219). The model may be used by policymakers to take action if there any disease or pest in the nursery or plantation.

Keywords—forecasting model; forest plantation; nursery; ARIMA; *Acacia mangium*, *Falcataria moluccana*

I. INTRODUCTION

The current trend in Indonesia is monoculture that has establishing of pure stands [1] and agroforestry system which combines agriculture and forest plantation. There is many problems with these plantation model, such as an out breaking of disease and pest because feed sources are available in abundance in a large area of plantation. Pest and diseases attacks have an impact on the procreation and quality of existing standing stocks, such as reducing a plantation growth and wood quality. Therefore, monitoring and identification of them has been needing to achieve of plantation health in agroforestry system. Previous works for coping with has been done by laboratory of forest health and protection Faculty of Forestry Universitas Gadjah Mada since 2013. It was starting from proposed model of expert system for identification pest and disease on *Falcataria moluccana* [2] and in 2017 is developing an extended approach model of expert system for various plantation based on mobile and website application [3]. In addition, related to these development model, we was published a model based on intelligent agent for monitoring of urban forest in 2014 [4] and developed an application for identification and calculation of size and intensity damage because of pest and diseases attack.

In this manuscript, we are proposing forecast model of pest and diseases on forest plantation. This model combined with the

expert system and the identification of wide and intensity pest and diseases attack model which was proposed by our previous works as an input. This model inspire from Edmonds et. al. [5] statements that the preliminary steps on forest disease management are disease identification, recognition, and forecasting. Furthermore we can considerate several important questions about management forest health, such as how can we predicting how fast outbreaks of diseases and its impact to plantation decease? How many decreasing of wood quantity if there is no disease treatments? How can we see and decide the plantation are using for conservation beneficial or the wood will be harvested? Accordingly, forecasting time series is an important problem which has been investigated by many researchers [6], especially in managing forest health and monitoring activities. Some researchers were conducted research on forecasting in several tree plantations, Watt et. al. [7] predict severity of *Dothistroma* needle blight disease and *Cyclaneusma* needle cast on *Pinus radiata* under future climate in New Zealand [8]. Moreover, Guzmán et. al. [9] was proposed the models for Bolivian *Pinus radiata* plantations by implementing a simulator for predict growing and future yields and to compare alternative management schedule. In addition, research finding for predicting *Acacia* hybrid (*A. auriculiformis* × *A. mangium*) productivity based on climate ranges and soil types have been done by Hung et al. [10], and Zhao et al. [11] have suggested the proposed model for prediction *Pinus sylvestris* L.var.monogolica Litvin Plantation. In addition, two methods of machine learning, Multivariate Adaptive Regression Splines (MARS) and M5 Model Trees (MT), have been reported by Rezaie-balf et.al. [12] to simulate the groundwater level (GWL) forecasting.

In this paper, forecast model combined with the expert system model, and the identification and calculation of damage size and the intensity of damage model. The expert system used for identification of pest and disease on plantation, model of identification used for collecting the data and calculating of a size and the intensity of plantation damage, and forecasting model will be used for seeing a disease growing without treatment. Proposed forecasting model will be applied on Powdery mildew diseases as data sample by using standard forecasting statistic model to test the data sample, such as Auto-Regressive, Moving Average, and ARIMA as a combined model. In the future research it will be evaluate and compare with several forecasting method to find the effective model.

This paper is organized as follows, section II discuss about method which is explaining a data collection, framework model, and others related method; In section III presents the results and

discussion to discuss the results and how the model tested by several forecasting method; and Section IV is giving a conclusion and the future works.

II. METHOD

A. Data collection

Data was collected from nursery of Forest Health and Protection Laboratory, Faculty of Forestry, Universitas Gadjah Mada which is taken by surveyors in the time series periods, approximately every three days since August 2016 until April 2017 (Fig. 1) containing a number of sick seedling over the time, and the trend of collected data shown in Fig. 2. The data sample used Powdery mildew disease which was inoculate into *Acacia mangium* seedling in the nursery.

```
DatetimeIndex(['2016-08-25', '2016-08-30', '2016-09-05', '2016-09-08',
              '2016-09-09', '2016-09-14', '2016-09-15', '2016-09-16',
              '2016-09-19', '2016-09-20', '2016-09-23', '2016-09-26',
              '2016-09-27', '2016-09-28', '2016-09-30', '2016-10-03',
              '2016-10-06', '2016-10-11', '2016-10-21', '2016-10-24',
              '2016-10-26', '2016-10-31', '2016-11-02', '2016-11-04',
              '2016-11-07', '2016-11-09', '2016-11-11', '2016-11-14',
              '2016-11-15', '2016-11-16', '2016-11-17', '2016-11-18',
              '2016-11-21', '2016-11-25', '2016-11-28', '2016-12-02',
              '2016-12-05', '2016-12-08', '2016-12-14', '2016-12-16',
              '2016-12-19', '2016-12-23', '2016-12-27', '2016-12-30',
              '2017-01-03', '2017-01-06', '2017-01-09', '2017-01-13',
              '2017-01-17', '2017-01-20', '2017-01-23', '2017-01-27',
              '2017-02-10', '2017-02-19', '2017-02-21', '2017-02-26',
              '2017-02-28', '2017-03-05', '2017-03-09', '2017-03-13',
              '2017-03-15', '2017-03-17', '2017-03-20', '2017-03-22',
              '2017-03-24', '2017-03-28', '2017-03-29', '2017-03-31',
              '2017-04-03'])
```

Fig. 1. Collecting date schedule of time series data

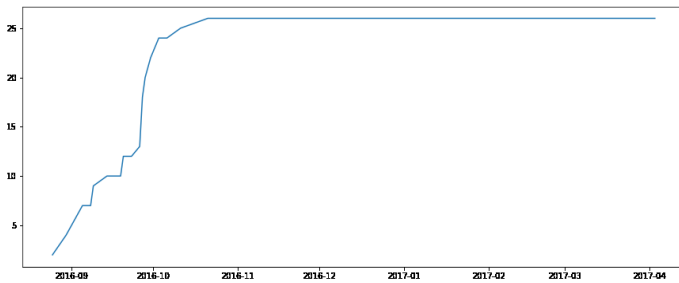


Fig. 2. Data collection of Powdery mildew

B. Framework model

Based on framework of integrated forest health and protection based software applications, we are developing three model starting from an expert system model, model of size and intensity damage, and the last model which explain in this paper is forecasting model, Fig. 3.

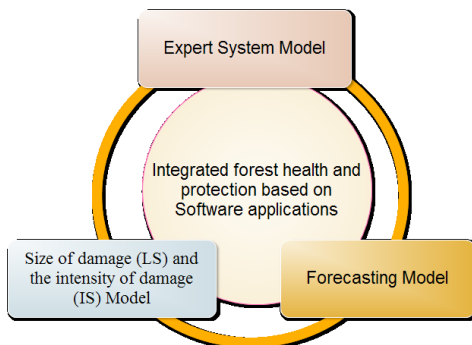


Fig. 3. Framework model integrated forest health and protection based on software applications

Expert system model was develop since 2013 when we were developing an initial system for identification pest and disease of *Falcataria moluccana* [2] and we are working up its model by evaluation the model and make some revisions to enhance utilization for several tree plantations. Size and intensity of damage model are used for collecting the data of forest health as preparing for forecasting model. Collecting approach has been explained in our previous work published by Pranolo and Widyastuti [4]. Therefore, in the last stage our forecasting model use data collection in the other two models.

C. Forecasting model

Fig. 4 shows a forecast model for forest plantation which divided into two models, namely forecasting model for seedling and forecasting model for tree plantation. Seedling forecast aims to forecast sick and dead seedling, meanwhile forecasting model of tree plantation has purposed for predicting damage size and intensity a plantation.

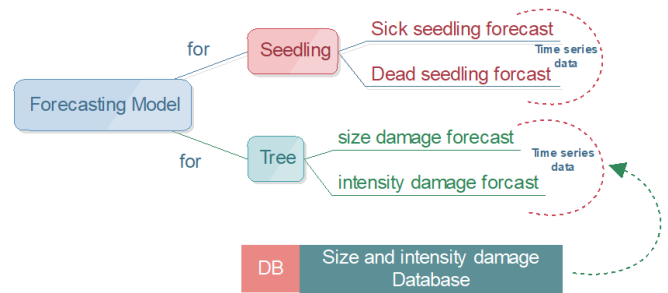


Fig. 4. Forecast model for forest plantation

In this paper will be focused on forecast testing for sick seedling by using standard forecast method.

D. Forecasting Method

1) Auto-regressive (AR) Model

An AR model is simply a linear regression of the current observation of the series against one or more prior observations of the series. The Autoregressive (AR) model time series has order p which is denoted by $AR(p)$ which is indicating that current observation depends on p past observations. The AR model is usually expressed as [13].

$$y^{(t)} = \sum_{i=1}^p \phi_i y^{(t-i)} + \epsilon_t \quad (1)$$

Where $y^{(t)}$ is the time series to be modeled, ϕ_i are the model coefficients, ϵ_t is the white noise which is independent from previous points, and order of AR model is given by parameter p .

2) Moving Average (MA) Model

In time series analysis, the moving-average (MA) model is a common approach for modeling univariate time series [14]. MA model defines that the variable of output depends on linear current and various past values of a stochastic (imperfectly predictable) terms. MA model order q can be defined as $MA(q)$ as.

$$y^{(t)} = \sum_{i=1}^q \phi_i + \epsilon_{t-1} \quad (1)$$

Where $y^{(t)}$ is the time series to be modeled, ϕ_i are the model coefficients, ϵ_t is the white noise which is independent from previous points, and order of MA model is given by parameter q .

3) Autoregressive integrated moving average (ARIMA)

An autoregressive integrated moving average (ARIMA) model is a generalization of an autoregressive moving average (ARMA) model [14][17]. The AR as part of ARIMA indicates which the evolving variable of interest is regressed on its own lagged (i.e., prior) values. The MA part defines that the regression error is actually a linear combination of error terms which values occurred contemporaneously and at various times in the past [14].

The ARIMA (p, d, q) model of a time series refers to the model which is stationary at level with p autoregressive term and q moving average term. It can be defined as [15]:

$$\Phi(B)(1 - B)^d Y_t = \theta(B)\epsilon_t, \tag{3}$$

where

$$\begin{aligned} \Phi(B) &= 1 - \Phi_1 B \\ &\quad \Phi_2 B^2 \dots \Phi_p B^p \text{ (autoregressive param)} \end{aligned} \tag{4}$$

$$\begin{aligned} \theta(B) &= 1 - \theta_1 B \\ &\quad \theta_2 B^2 \dots \theta_q B^q \text{ (moving average param)} \end{aligned} \tag{5}$$

ϵ_t is the white noise term, and B is the backshift operator.

III. RESULTS AND DISCUSSION

A. Stationary Test Results

The stationer data can be seen if there is no growing and declining which signed by the data pattern is on the equilibrium around the constant value and mean of variance is constant for a certain time, Fig. 4. To check the stationary, we use plotting rolling statistics by using moving average and dickey fuller test [16] which is one of the statistical test that using null hypothesis to determine non-stationary data. Fig. 5 shows that standard deviation is small and the rolling mean increase slightly on the first three months, accordingly it shows an evidence that the data is not stationary.

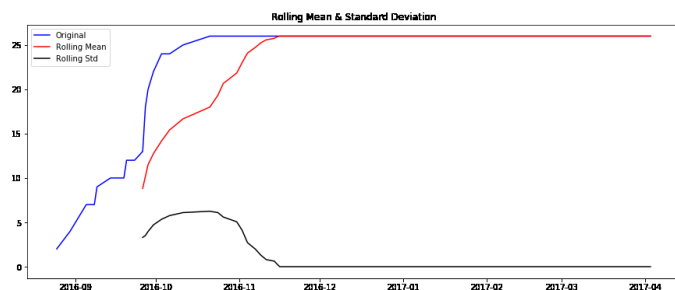


Fig. 5. Stationary test results

The process moving average has been done to take average of 'k' consecutive values depending on the frequency of time series. The results of Dickey-Fuller test shown test statistics value is -2.041262, p-value 0.268779, 10.000000 Lags Used, 58.000000 number of observations used, critical value 5% is -2.912837, 1% is -3.548494, and 10% is -2.594129. It can be represented that the 2nd time series stationary test has more good than 1st test because the statistics test smaller than 5% of critical values which can be claimed that stationary has 95% confidence, Fig. 6.

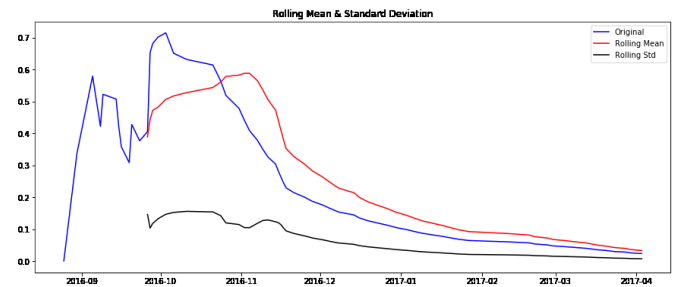


Fig. 6. Stationary test results

Besides from the time series plot, a stationary can be seen from the graph of Autocorrelation Function (AF) of the data. Fig 7(a). Shows that AF decrease dramatically on lag 2 close to 0 value, we can say that the data is stationer.

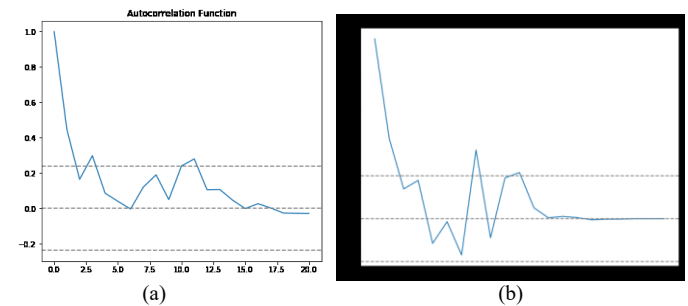


Fig. 7. Autocorrelation and partial autocorrelation function

AF and partial autocorrelation function (PAF), Fig 7(b) are indicating which values of past series are most useful in predicting future values. AF and PAF can also measure of association between current and past series values.

Non-stationary data can be seen if there have elements of trends, increase or decrease is depend on increasing period of time. For the example, non-stationary data sometimes has a trend indicated by AF values are high in the initial lag and decrease gradually in the next lag.

B. Removing trend and seasonality

There is has a technique for removing trend and seasonality by using differencing and decomposition. In this case we are using decomposition technique, Fig. 8 with modelling of trend and seasonality and removing them from the model. Fig. 8 describes that trend, seasonality, and residuals are separated from the original data. The results of Dickey-Fuller test statistics is still significantly lower than one 1% critical value.

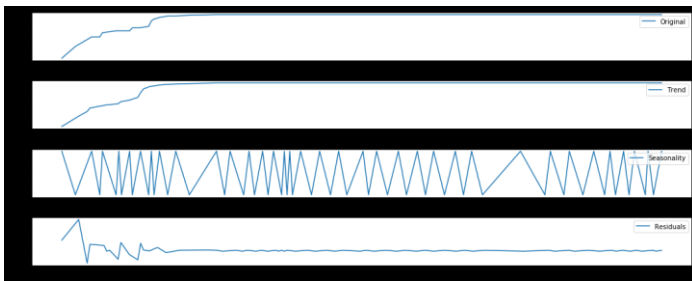


Fig. 8. Decomposition technique

C. AR, MA, and ARIMA Model

The Autoregressive Integrated Moving Average (ARIMA) model is nonstationary ARMA model which has diverted into stationary model. The predictors depend on the parameters (p,d,q) of the ARIMA model which Number of AR (Auto-Regressive) terms (p) are just lags of dependent variable, while Number of MA terms (q) are lagged forecast errors in prediction equation. In Fig. 11 shows that RSS for ARIMA model is 0.7791, it is better than RSS which is resulting by AR, Fig. 9 and MA, Fig. 10 models.

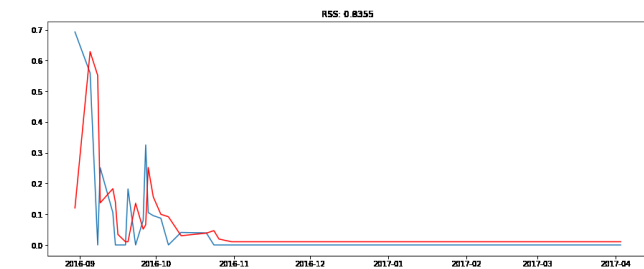


Fig. 9. AR Model

Fig. 9 shows that RSS for AR model is 0.8355 and MA model is 0.8013, Fig. 10.

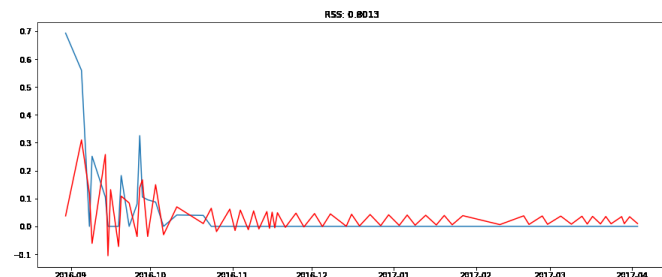


Fig. 10. MA Model

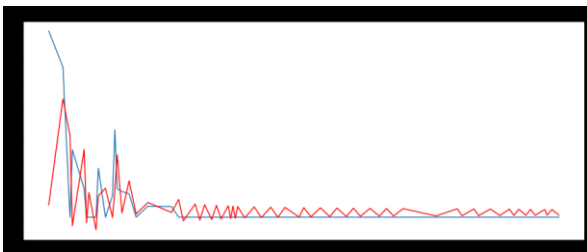


Fig. 11. ARIMA Model

To measure the accuracy of forecast model, we are using root mean square error (RMSE). The best RMSE value for ARIMA is 7.9291 shown in Fig. 12.

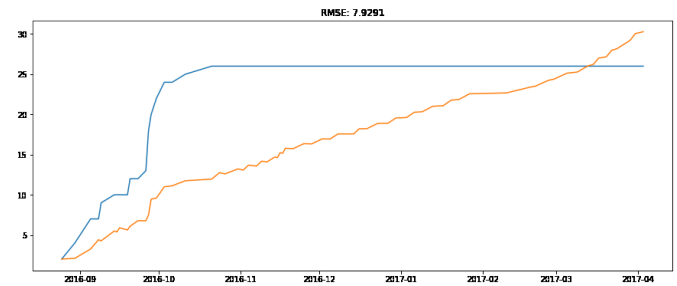


Fig. 12. RMSE Value

IV. CONCLUSION

In this paper we are proposed forest plantation forecast model which divided into two models, namely forecasting model for seedling and forecasting model for tree plantation. Seedling forecast aims to forecast sick and dead seedling, meanwhile forecasting model of tree plantation has purposed for predicting damage size and intensity a plantation. We are using order (2,1,0) for AR, MA (0,1,2), and combined method ARIMA (2,1,1) which has better RSS value (0.6219). The model may be used by policymakers to take action if there any disease or pest in the nursery or plantation.

ACKNOWLEDGMENT

The authors wish to acknowledge the financial support provided by Directorate General Research Technology and Higher Education (RistekDikti) Republic of Indonesia through Penelitian Terapan Unggulan Perguruan Tinggi (PTUPT) grant 2017.

REFERENCES

- [1] A. Mamle and R. Jolanda, "Disease of plantation forestry trees in shouthern ghana," *Int. J. Phytopathol.*, vol. 4, no. 1, pp. 5–13, 2015.
- [2] A. Pranolo, "Expert System for Disease Identification on Sengon (Falcataria moluccana)," Universitas Gadjah Mada, 2013.
- [3] A. Pranolo, S. M. Widyastuti, and Azhari, "Expert system model for identification pests and diseases of forest tree plantations," *Int. J. Adv. Soft Comput. its Appl.*, vol. 9, no. 2, pp. 204–216, 2017.
- [4] A. Pranolo and S. M. Widyastuti, "Simple additive weighting method on intelligent agent for urban forest health monitoring," in *2014 International Conference on Computer, Control, Informatics and Its Applications (IC3INA)*, 2014, pp. 132–135.
- [5] R. L. Edmonds, J. K. Agee, and R. I. Gara, *Forest Health and Protection*. McGraw-Hill Higher Education, 2000.
- [6] J. L. Salmeron and W. Froelich, "Dynamic optimization of fuzzy cognitive maps for time series forecasting," *Knowledge-Based Syst.*, vol. 105, pp. 29–37, 2016.
- [7] M. S. Watt, D. J. Palmer, and L. S. Bulman, "Predicting the severity of Dothistroma needle blight on Pinus radiata under future climate in New Zealand," *New Zeal. J. For. Sci.*, vol. 41, pp. 207–215, 2011.
- [8] M. S. Watt, C. A. Rolando, D. J. Palmer, and L. S. Bulman, "Predicting the severity of Cyclaneusma minus on Pinus radiata under current climate in New Zealand," *Can. J. For. Res.*, vol. 42, no. 4, pp. 667–674, Apr. 2012.

- [9] G. Guzmán, T. Pukkala, M. Palahí, and S. de-Miguel, "Predicting the growth and yield of *Pinus radiata* in Bolivia," *Ann. For. Sci.*, vol. 69, no. 3, pp. 335–343, Dec. 2011.
- [10] T. T. Hung, A. C. Almeida, A. Eyles, and C. Mohammed, "Predicting productivity of Acacia hybrid plantations for a range of climates and soils in Vietnam," *For. Ecol. Manage.*, vol. 367, pp. 97–111, May 2016.
- [11] L. J. Zhao, F. R. Li, and W. W. Jia, "Prediction Model for *Pinus sylvestris* *L.var.monogolica* Litvin Plantation," *Appl. Mech. Mater.*, vol. 253–255, no. PART 1, pp. 1055–1062, Dec. 2012.
- [12] M. Rezaie-balf, S. R. Naganna, A. Ghaemi, and P. C. Deka, "Wavelet coupled MARS and M5 model tree approaches for groundwater level forecasting," *J. Hydrol.*, 2017.
- [13] Y. Zhang, B. Liu, X. Ji, and D. Huang, "Classification of EEG signals based on autoregressive model and wavelet packet decomposition," *Neural Process. Lett.*, vol. 45, no. 2, pp. 365–378, 2017.
- [14] N. A. Bakar and S. Rosbi, "Data Clustering using Autoregressive Integrated Moving Average (ARIMA) model for Islamic Country Currency: An Econometrics method for Islamic Financial Engineering," *Int. J. Eng. Sci.*, vol. 6, no. 6, pp. 22–31, 2017.
- [15] J. Unnikrishnan and K. K. Suresh, "Modelling the Impact of Government Policies on Import on Domestic Price of Indian Gold Using ARIMA Intervention Method," *Int. J. Math. Math. Sci.*, vol. 2016, 2016.
- [16] A. Jain, "Complete guide to create a Time Series Forecast (with Codes in Python)," 2016. [Online]. Available: <https://www.analyticsvidhya.com/blog/2016/02/time-series-forecasting-codes-python/>. [Accessed: 04-Sep-2017].
- [17] Haviluddin and A. Jawahir, "Comparing of ARIMA and RBFNN for short-term forecasting," *International Journal of Advances on Intelligent Informatics (IJAIN)*, vol. 1, (1), March 2015, pp. 15-22, 2015.