

Modeling the Dynamics of Forest Fires: A Vector Autoregressive Approach Across Three Fire Classifications

Nur'ainul Miftahul Huda1, Nurfitri Imro'ah²

¹Mathematics Department, Universitas Tanjungpura Pontianak, Indonesia ²Statistics Department, Universitas Tanjungpura Pontianak, Indonesia nurainul@fmipa.untan.ac.id

A. INTRODUCTION

Many empirical research uses time series data, comprising observations from several variables. Within the context of a study on sales performance, for instance, the variables that could be considered include sales volume, prices, sales force, and advertising expenditures. An input-output relationship between one output variable and one or more input variables is the subject of investigation in the transfer function model. This interaction between variables is referred to as an input-dependent relationship. Within regression time series models, the primary focus is the relationship between a dependent variable and a collection of independent variables (Wei, 2006). On the other hand, transfer function and regression models might only be suitable for use in some applications. For this study, the Vector Autoregressive (VAR) model, which belongs to a more general category of vector time series models, is utilized to

—————————— ——————————

characterize the relationships between several time series variables (Cryer & Chan, 2008). Some research about VAR model is done by many researchers (Apergis & Apergis, 2021; Box & Jankins, 1976; Fruet Dias & Kapetanios, 2017; Kaur et al., 2023a, 2023b; Khan et al., 2020; Meimela et al., 2021; Ramyar & Kianfar, 2019; Reinsel, 1993; Schaffer et al., 2021; Singh et al., 2020; Tsay, 2014; Xu et al., 2021). This research applied the VAR model to forest fire data from 2013 – 2024.

Over the past few years, forest fires have emerged as a significant environmental problem, and Kalimantan Island, which is in Indonesia and is a part of Borneo, has been confronted with this issue regularly. A few reasons, including illicit logging, land removal for agricultural purposes, and dry weather conditions, frequently exacerbate these fires. The dense tropical rainforests of Kalimantan, renowned for their abundant biodiversity and one-of-a-kind ecosystems, are especially susceptible to these fires, which cause severe ecological damage, contribute to air pollution, and threaten the lives of the populations that live there. To protect the natural heritage of Kalimantan Island and address broader environmental challenges in the region, it is essential to have a solid understanding of the factors that cause these fires, their effects, and the measures being taken to mitigate them.

Pulang Pisau is well-known for its vast woods, which are essential for the preservation of biodiversity and the populations that are dependent on the resources provided by the forest. Land clearance practices, such as slash-and-burn agriculture, in which farmers burn vegetation to prepare the ground for planting crops, are frequently the cause of these fires. Many other factors can also contribute to the occurrence of these fires. On the other hand, these flames have the potential to rapidly expand out of control, particularly during dry seasons that are made worse by climate change. Because peat may smoulder underground for extended periods, when it is lit, it releases a significant amount of smoke and greenhouse gasses. As a result, the peatlands in Pulang Pisau are especially prone to fires. Some research about forest fires are done by many researchers (Alisjahbana & Busch, 2017; Carmenta et al., 2017a, 2017b; Carta et al., 2023; Clarke et al., 2022a; Edwards & Heiduk, 2015; Ghorbanzadeh et al., 2019a; Muhammad et al., 2019).

In the context of environmental risk management and mitigation, forest fire monitoring is an essential component, particularly in places that are susceptible to environmental hazards, such as Pulang Pisau. Given the effects of climate change and the growing complexity of human activities, it is essential to have a solid understanding of the patterns and trends of forest fires. In order to accomplish this, this research makes use of historical data on forest fires that have been classified into three different levels of intensity: high, medium, and low. This classification was accomplished by employing a Vector Autoregression (VAR) model technique. Using this methodology, the purpose of the study is to provide detailed forecasts of the number of forest fires that are expected to occur in Pulang Pisau over the following three periods. This will consequently provide a solid foundation for more effective preparation and reaction in the face of future disasters.

This article provides an explanation of four primary sections, the first of which is an introduction concerning the VAR model and the conditions of forest fires. Further, the second section demonstrates the theoretical foundation that is associated with the method that was utilized. A comprehensive explanation of how the VAR approach is utilized to assess forest fire

concerns is presented in the third section of this article. The conclusion and discussion are included in the final section.

B. METHODS

A Vector Autoregressive (VAR) model is a type of multivariate time series model that describes the joint dynamics of numerous time series variables. This form of model is also known as a VAR model. It assumes that every variable in the system is modelled as a linear function of its own lagged values and the lagged values of every other variable in the system. The general vector AR(p) process (Meimela et al., 2021; Wei, 2006)

$$
(I - \Phi_1 B - \cdots - \Phi_p B^p)Z_t = e_t
$$

$$
Z_t = \Phi_1 Z_{t-1} + \cdots + \Phi_p Z_{t-p} + e_t
$$

where $\Phi_1, ..., \Phi_n$ are matrix of parameter in VAR(p). Consider the two-dimensional VAR(1) process,

$$
(I - \Phi_1 B)Z_t = e_t
$$

$$
Z_t = \Phi_1 Z_{t-1} + e_t
$$

where,

or

$$
\Phi_1 = \begin{bmatrix} \phi_{11} & \phi_{12} \\ \phi_{21} & \phi_{22} \end{bmatrix}; \ \Sigma = E(e_t e_t') = \begin{bmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{21} & \sigma_{22} \end{bmatrix}
$$

and

$$
\Gamma(0) = E(\mathbf{Z}_t \mathbf{Z}_t') = \begin{bmatrix} \gamma_{11}(0) & \gamma_{12}(0) \\ \gamma_{21}(0) & \gamma_{22}(0) \end{bmatrix}
$$

The following are the main steps in VAR model analysis (Box & Jankins, 1976; Reinsel, 1993; Tsay, 2014; Wei, 2006),

- 1. Model Identification. In principle, identification of VAR models is similar to the identification of univariate time series models. For a given observed vector time series $Z_1, Z_2, ..., Z_n$, the identification uses the pattern of its sample correlation and partial correlation matrices after proper transformations are applied to reduce a nonstationary series to stationary series.
- 2. Parameter Estimation. Determination the values of the coefficients $\Phi_1 = \begin{bmatrix} \phi_{11} & \phi_{12} \\ \phi_{21} & \phi_{22} \end{bmatrix}$ $\begin{bmatrix} \varphi_{11} & \varphi_{12} \\ \varphi_{21} & \varphi_{22} \end{bmatrix}$ by methods such as the Least Squares estimation. The software used in parameter estimation is *R*.
- 3. Diagnostic Checking. Examination of residuals for autocorrelation and normality assumptions is an appropriate method to determine whether the model is adequate.

C. RESULT AND DISCUSSION

1. Descriptive Statistics

The data utilized is secondary data, namely data on the number of forest fires in Pulang Pisau (Central Kalimantan). These fires are classified into three categories: high, medium, and low. As shown in Figure 1, the map shows the number of hotspots found in Kalimantan Island.

Figure 1. Map of the Number of Hotspots on Kalimantan Island

On the island of Kalimantan, Figure 1 is a clustered display of the total number of forest fire cases that occurred from 2013 to 2024 in each district. In total, there are 51 districts and cities spread throughout four provinces: East Kalimantan -EK- (ten), West Kalimantan -WK- (fourteen), Central Kalimantan -CK- (fourteen), and South Kalimantan -SK- (thirteen). According to the clustering findings using a K-means cluster using three classifications (high, moderate, and low), it is possible to determine which districts and cities fall into the high group (that is, districts and cities with the highest number of cases). Detailed information regarding a list of districts and cities according to classification may be found in Table 1.

Pulang Pisau district, which is in Central Kalimantan, is the district or city that has the greatest total number of forest fire instances out of all the districts and cities in the state, according to Table 1. March 2024 to January 2013 is the period that is being used. Following this, the data is separated into two distinct categories: the training data, which accounts for 97% of the total data and consists of 129 observations, and the testing data, which accounts for 3% of the total data and consists of 5 observations. Training data is used to evaluate how well the model estimates data, whilst testing data is used to evaluate how well the model predicts data. Both types of data are used to evaluate the effectiveness of the model. Figure 2 depicts a plot of time series data that is divided into three categories: the top plot represents the high forest fire category (represented by the red line), the next plot in the middle and bottom plots represent the medium forest fire category (represented by the green line), and the bottom plot represents the low forest fire category (represented by the yellow line). Figure 2 makes it very evident that there is an obvious boost at periods, namely during September and October. The rise in the number of forest fires that occurred in Central Kalimantan during September and October can be attributed to several factors that have an impact on the natural conditions of the region, including the following (Alisjahbana & Busch, 2017; Clarke et al., 2022a):

- a. Central Kalimantan experiences its longest dry season throughout the months of September and October, which is often the driest time of the year.
- b. The Practice of Burning Land. In Central Kalimantan, the traditional practice of burning land to clear land for agricultural or forestry purposes is still widely practiced.
- c. Peatlands. Central Kalimantan is home to a significant number of peatlands, which are prone to being consumed by fire. Peat is an organic material that, as it dries out, becomes highly flammable.

September and October are the most susceptible months for forest fires in Central Kalimantan due to the confluence of the circumstances described above, as shown in Figure 2.

Figure 2. Plot of Number of the Hotspots in Pulang Pisau (Central Kalimantan)

A breakdown of descriptive statistics for each category is presented in Table 2. There is no data in any category that is less than zero. This is because there were no reports of forest fires in Pulang Pisau during specific months, particularly during the wet season. On the other hand, the highest data in the high categorization reveals that approximately two thousand cases were discovered in just one month, specifically in September 2019. Even more concerning is that the medium categorization had the highest number of instances discovered in October 2023, which was 4,297. While this was happening, September 2015 saw the highest number of cases in the low classification, specifically 524 cases. The average number of monthly cases for the High, Medium, and Low classifications is 58, 148, and 19, respectively. According to estimates, approximately 225 forest fires occur every month in Pulang Pisau (Central Kalimantan). Considering the magnitude of forest fires' impact, this is not a tiny number by any stretch of what can be expected.

Table 2. Descriptive Statistics

2. Vector Autoregressive Model

a. Order Identification

The order that was selected is order 4, or VAR(4), by using Rstudio to compare the AIC values of the potential orders (limited to order 10). This is how the VAR(4) model is formed.

$$
\begin{bmatrix} H_t \\ M_t \\ L_t \end{bmatrix} = \begin{bmatrix} \phi_0^{(H)} \\ \phi_0^{(M)} \\ \phi_0^{(L)} \end{bmatrix} + \begin{bmatrix} \phi_{11}^{(1)} & \phi_{12}^{(1)} & \phi_{13}^{(1)} \\ \phi_{21}^{(1)} & \phi_{22}^{(1)} & \phi_{23}^{(1)} \\ \phi_{31}^{(1)} & \phi_{32}^{(1)} & \phi_{33}^{(1)} \end{bmatrix} \begin{bmatrix} H_{t-1} \\ M_{t-1} \\ L_{t-1} \end{bmatrix} + \dots + \begin{bmatrix} \phi_{11}^{(4)} & \phi_{12}^{(4)} & \phi_{13}^{(4)} \\ \phi_{21}^{(4)} & \phi_{22}^{(4)} & \phi_{23}^{(4)} \\ \phi_{31}^{(4)} & \phi_{32}^{(4)} & \phi_{33}^{(4)} \end{bmatrix} \begin{bmatrix} H_{t-4} \\ M_{t-4} \\ L_{t-4} \end{bmatrix} + \begin{bmatrix} e_t^{(H)} \\ e_t^{(M)} \\ e_t^{(L)} \end{bmatrix}
$$

b. Parameter Estimation

Following the completion of the preceding stages, the VAR model order chosen is VAR(4). The estimated parameters are presented in Table 3, which were generated using the least squares method.

The VAR(4) model equation for each classification is displayed in Table 3 below, along with the statistical results of the F test (to determine the significance of the influence of the independent variable on the dependent variable) and the R-squared value (to determine how much influence there is). These results are based on the results of the performed parameter estimates.

c. Diagnostic Checking

Diagnostic tests are performed on the VAR(4) model, including residual autocorrelation and residual independence tests (Table 4). It can be concluded that the VAR(4) model is adequate for estimating the number of forest fire incidents in Pulang Pisau because each categorization demonstrates that each VAR(4) structural model satisfies these two requirements.

Table 4. Diagnostic Checking

Once the diagnostic test has been completed, the next step is to perform estimates using the VAR(4) model on the training data, as demonstrated in Figure 3. The black dots show the fitted values based on the VAR(4) model, whereas the other coloured dots have the same meaning as in Figure 2. From Figure 3, the VAR(4) model can provide a satisfactory representation of the training data. Taking into consideration the MAPE value of 29.77%, this is evident. After that, the VAR(4) model is utilized on the data just tested. A MAPE value of 38.41% was found to be the outcome obtained. Consequently, this indicates that the VAR(4) model is far more effective when estimating data than predicting forest fires.

Figure 3. Fitted Values vs Observations Data based on VAR(4) model

Following this, forecasts are generated for the subsequent three time periods by employing the VAR(4) model, which is depicted in Figure 4.

Forecasting Values for 3 Period Ahead

Figure 4. Forecasting Values based on VAR(4) model

Some studies have discussed the VAR method, but these are distinct case applications when examining the previous research (Fall et al., 2020; Rusyana et al., 2020; Suhartono et al., 2018). Similarly, when considering the case of forest fires, certain studies employ distinct methodologies to analyze the case (Clarke et al., 2022b; Ghorbanzadeh et al., 2019b; Rossi & Becker, 2019).

D. CONCLUSION AND SUGGESTIONS

Using the VAR(4) model, it is possible to accurately estimate the situation regarding the forest fires in Pulang Pisau. The interpretation of the VAR(4) model is that the cases of forest fires that happened in Pulang Pisau for each classification were influenced up to the previous fourth period. This is the case for all of the classifications. For example, according to the model that was obtained in the previous discussion, the number of forest fires that have occurred in the Medium classification is currently influenced by the number of forest fires that have occurred in the Medium, Low, and High classifications in the preceding one month, two months, three months, and also four months prior to the current period. Each variable's parameter coefficients provide evidence of this influence, which can be observed. As far as the High and Low classes are concerned, the same interpretation applies. The classification at that time (at time t) is collectively influenced by each classification that occurred in the preceding one to four periods, according to the F test, if we look at the partial influence. In other words, increasing the number of forest fires that fall under the Medium classification could be influenced not only by the number of forest fires that occurred in the Medium classification during the four periods that came before but also by the other categories that occurred during those same four periods. When seen from the standpoint of the Low classification, the number of forest fire cases that have this classification significantly impacts the process of forest fire cases being elevated to the medium or high classification levels. Furthermore, the same holds for other classes. This is an attempt to put out fires and prevent fires from occurring in the first place so that forest fires can be put out. The results of the predictions made for the subsequent three periods, provided before, can serve as the basis for these thoughts. There was a large increase from one case to fifteen cases from May to June, and it is vital to be cautious with Medium instances during this period (without neglecting other classes).

REFERENCES

- Alisjahbana, A. S., & Busch, J. M. (2017). Forestry, Forest Fires, and Climate Change in Indonesia. *Bulletin of Indonesian Economic Studies*, *53*(2), 111–136. https://doi.org/10.1080/00074918.2017.1365404
- Apergis, E., & Apergis, N. (2021). The impact of COVID-19 on economic growth: evidence from a Bayesian Panel Vector Autoregressive (BPVAR) model. *Applied Economics*, *53*(58), 6739–6751. https://doi.org/10.1080/00036846.2021.1946479
- Box, G. E. P., & Jankins, G. M. (1976). *Time Series Analysis Forecasting and Control*. Wiley.
- Carmenta, R., Zabala, A., Daeli, W., & Phelps, J. (2017a). Perceptions across scales of governance and the Indonesian peatland fires. *Global Environmental Change*, *46*, 50–59. https://doi.org/10.1016/j.gloenvcha.2017.08.001
- Carmenta, R., Zabala, A., Daeli, W., & Phelps, J. (2017b). Perceptions across scales of governance and the Indonesian peatland fires. *Global Environmental Change*, *46*, 50–59. https://doi.org/10.1016/j.gloenvcha.2017.08.001
- Carta, F., Zidda, C., Putzu, M., Loru, D., Anedda, M., & Giusto, D. (2023). Advancements in Forest Fire Prevention: A Comprehensive Survey. *Sensors*, *23*(14), 6635. https://doi.org/10.3390/s23146635
- Clarke, H., Nolan, R. H., De Dios, V. R., Bradstock, R., Griebel, A., Khanal, S., & Boer, M. M. (2022a). Forest fire threatens global carbon sinks and population centres under rising atmospheric water demand. *Nature Communications*, *13*(1), 7161. https://doi.org/10.1038/s41467-022-34966- 3
- Clarke, H., Nolan, R. H., De Dios, V. R., Bradstock, R., Griebel, A., Khanal, S., & Boer, M. M. (2022b). Forest fire threatens global carbon sinks and population centres under rising atmospheric water demand. *Nature Communications*, *13*(1), 7161. https://doi.org/10.1038/s41467-022-34966- 3
- Cryer, D. J., & Chan, K.-S. (2008). *Time Series Analysis: With Application in R*. Springer Science and Bussiness Media.
- Edwards, S. A., & Heiduk, F. (2015). Hazy Days: Forest Fires and the Politics of Environmental Security in Indonesia. *Journal of Current Southeast Asian Affairs*, *34*(3), 65–94. https://doi.org/10.1177/186810341503400303
- Fall, S., N'Guessan, A., Iraci, F., & Koutouan, A. (2020). *Forecasting the French Personal Services Sector Wage Bill: A VARIMA Approach* (pp. 119–134). https://doi.org/10.1007/978-3-030-13697- 0_9
- Fruet Dias, G., & Kapetanios, G. (2017). Supplement To "Estimation and Forecasting in Vector Autoregressive Moving Average Models for Rich Datasets." *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.2830838
- Ghorbanzadeh, O., Blaschke, T., Gholamnia, K., & Aryal, J. (2019a). Forest Fire Susceptibility and Risk Mapping Using Social/Infrastructural Vulnerability and Environmental Variables. *Fire*, *2*(3), 50. https://doi.org/10.3390/fire2030050
- Ghorbanzadeh, O., Blaschke, T., Gholamnia, K., & Aryal, J. (2019b). Forest Fire Susceptibility and Risk Mapping Using Social/Infrastructural Vulnerability and Environmental Variables. *Fire*, *2*(3), 50. https://doi.org/10.3390/fire2030050
- Kaur, J., Parmar, K. S., & Singh, S. (2023a). Autoregressive models in environmental forecasting time series: a theoretical and application review. *Environmental Science and Pollution Research*, *30*(8), 19617–19641. https://doi.org/10.1007/s11356-023-25148-9
- Kaur, J., Parmar, K. S., & Singh, S. (2023b). Autoregressive models in environmental forecasting time series: a theoretical and application review. *Environmental Science and Pollution Research*, *30*(8), 19617–19641. https://doi.org/10.1007/s11356-023-25148-9
- Khan, F., Saeed, A., & Ali, S. (2020). Modelling and forecasting of new cases, deaths and recover cases of COVID-19 by using Vector Autoregressive model in Pakistan. *Chaos, Solitons & Fractals*, *140*, 110189. https://doi.org/10.1016/j.chaos.2020.110189
- Meimela, A., Lestari, S. S. S., Mahdy, I. F., Toharudin, T., & Ruchjana, B. N. (2021). Modeling of Covid-19 in Indonesia using Vector Autoregressive Integrated Moving Average. *Journal of Physics: Conference Series*, *1722*(1), 012079. https://doi.org/10.1088/1742-6596/1722/1/012079
- Muhammad, K., Khan, S., Elhoseny, M., Hassan Ahmed, S., & Wook Baik, S. (2019). Efficient Fire Detection for Uncertain Surveillance Environment. *IEEE Transactions on Industrial Informatics*, *15*(5), 3113–3122. https://doi.org/10.1109/TII.2019.2897594
- Ramyar, S., & Kianfar, F. (2019). Forecasting Crude Oil Prices: A Comparison Between Artificial Neural Networks and Vector Autoregressive Models. *Computational Economics*, *53*(2), 743–761. https://doi.org/10.1007/s10614-017-9764-7
- Reinsel, G. C. (1993). *Vector ARMA Time Series Models and Forecasting* (pp. 21–51). https://doi.org/10.1007/978-1-4684-0198-1_2
- Rossi, F., & Becker, G. (2019). Creating forest management units with Hot Spot Analysis (Getis-Ord Gi*) over a forest affected by mixed-severity fires. *Australian Forestry*, *82*(4), 166–175. https://doi.org/10.1080/00049158.2019.1678714
- Rusyana, A., Tatsara, N., Balqis, R., & Rahmi, S. (2020). Application of Clustering and VARIMA for Rainfall Prediction. *IOP Conference Series: Materials Science and Engineering*, *796*(1), 012063. https://doi.org/10.1088/1757-899X/796/1/012063
- Schaffer, A. L., Dobbins, T. A., & Pearson, S.-A. (2021). Interrupted time series analysis using autoregressive integrated moving average (ARIMA) models: a guide for evaluating large-scale health interventions. *BMC Medical Research Methodology*, *21*(1), 58. https://doi.org/10.1186/s12874-021-01235-8
- Singh, S., Parmar, K. S., Kumar, J., & Makkhan, S. J. S. (2020). Development of new hybrid model of discrete wavelet decomposition and autoregressive integrated moving average (ARIMA) models in application to one month forecast the casualties cases of COVID-19. *Chaos, Solitons & Fractals*, *135*, 109866. https://doi.org/10.1016/j.chaos.2020.109866
- Suhartono, S., Prastyo, D. D., Kuswanto, H., & Lee, M. H. (2018). Comparison between VAR, GSTAR, FFNN-VAR and FFNN-GSTAR Models for Forecasting Oil Production. *MATEMATIKA*, *34*(1), 103–111. https://doi.org/10.11113/matematika.v34.n1.1040
- Tsay, R. S. (2014). *Multivariate Time Series Analysis with R and Financial Applications.* Wiley.
- Wei, W. W. S. (2006). *Time Series Analysis Univariate and Multivariate Methods*(2nd ed.). Addison Wesley.
- Xu, H., Ding, F., & Yang, E. (2021). Three‐stage multi‐innovation parameter estimation for an exponential autoregressive time‐series model with moving average noise by using the data filtering technique. *International Journal of Robust and Nonlinear Control*, *31*(1), 166–184. https://doi.org/10.1002/rnc.5267