Forecasting and Mapping Coffee Borer Beetle Attacks Using GSTAR-SUR Kriging and GSTARX-SUR Kriging Models

Henny Pramoedyo[1]*, Arif Ashari[2], and Alfi Fadliana[3]

[1-3]Department of Statistics, Faculty of Mathematics and Natural Sciences, Brawijaya University
Jln. Veteran, Malang 65145, Indonesia
[1]hennyp@ub.ac.id; [2]arif.7684@gmail.com; [3]alfifadliana@gmail.com

Abstract - The research aimed to use Generalized Space Time Autoregressive (GSTAR) and GSTARX modeling with the Seemingly Unrelated Regression (SUR) approach and combine them with the Kriging interpolation technique in an unobserved location. The case study was coffee borer beetle forecasting in Probolinggo Regency, East Java, Indonesia, with Watupanjang Village as the unobserved location. The results show that GSTAR-SUR Kriging and GSTARX-SUR Kriging models can predict coffee borer beetle attacks in unobserved areas with high accuracy. It is indicated by the Mean Absolute Percentage Error (MAPE) values of less than 10%. The addition of exogenous variables (rainfall) into the model is proven to improve the accuracy of the model. The Root-Mean-Square Error (RMSE) value of the GSTARX-SUR Kriging model is smaller than the GSTAR-SUR Kriging model. The structure of the model produced from the research, GSTARX-SUR (1,[1,12])(10,0,0), can be used as a reference in modeling coffee borer beetle attacks in other regencies. Map of forecasting coffee borer beetle attack shows that the spread of coffee borer beetle attack is spatial clustering with the attack center located in the eastern region of Probolinggo Regency.

Keywords: coffee borer beetle, Generalized Space Time Autoregressive (GSTAR), GSTARX, Seemingly Unrelated Regression (SUR), Kriging

I. INTRODUCTION

Generalized Space Time Autoregressive (GSTAR) is a multivariate time series data model that involves location aspects. The GSTAR model is a natural generalization of Space Time Autoregressive (STAR) models, allowing the autoregressive parameters to vary per location. Hence, the GSTAR model is applicable to the heterogeneous characteristic of sample locations (Ruchjana, Borovkova, & Lopuhaa, 2012). The parameter estimation of the GSTAR model with Ordinary Least Square (OLS) has weaknesses

when inter-location residuals are correlated. To overcome this issue, Seemingly Unrelated Regression (SUR) approach is done using Generalized Least Square (GLS) as a parameter estimation method for the model (Iriany, Suhariningsih, Ruchjana, & Setiawan, 2013). GSTAR model with the SUR approach is known as GSTAR-SUR.

The GSTAR model will undoubtedly be more informative and useful by adding exogenous variables to the model. Because exogenous variables are also affected by time, the model can use the transfer function approach. Modeling by entering exogenous variables into the GSTAR model is known as GSTARX.

Suhartono, Wahyuningrum, Setiawan, and Akbar (2016) developed the GSTARX-GLS model in forecasting the inflation rate in four major cities in East Java. They used the increase in fuel oil as a non-metric exogenous variable. The results showed that GSTARX was better than the Vector Autoregressive Integrated Moving Average with Exogenous Variable (VARIMAX) model. Moreover, Astuti, Ruchjana, and Soemartini (2017) applied GSTARX model to predict an export volume of Crude Palm Oil (CPO) in several locations on the island of Sumatera, in which X is the international CPO prices. CPO export volume in one area was affected by the CPO export volume in the past in the same location, the volume of CPO exports in the past in other areas, and international CPO prices.

Meanwhile, Andayani, Sumertajaya, Ruchjana, and Aidi (2017) compared the Generalized Space Time Autoregressive Integrated Moving Average (GSTARIMA) and Generalized Space Time Autoregressive Integrated Moving Average with Exogenous Variable (GSTARIMAX) models in approaching rice price data in six provinces in Java. In the research, the exogenous variables used were metric data, namely milled dry grain price. The results showed that the GSTARIMAX was better than the GSTARIMA model. The three previous researchers prove that adding exogenous variables into the GSTAR model can improve forecasting accuracy.

GSTAR model normally can only predict an event in the future in locations where data are indeed used to build the model. The case that often happens is that a location does

not have data or complete data. Hence, several alternatives can be done, one of which is to combine the GSTAR model with interpolation techniques. Research on this matter has been carried out by Abdullah *et al.* (2018). They combined GSTAR with Kriging interpolation, known as GSTAR-Kriging model. In this previous research, exogenous variables are not added to the model.

According to Ruchjana *et al.* (2012), GSTAR model with autoregressive order (p) and spatial order (λ) can be written as follows:

$$z_{(t)} = \sum_{k=1}^{p} \sum_{l=0}^{\lambda_k} \phi_{kl} W^{(l)} z_{(t-k)} + \varepsilon_{(t)} \qquad (1)$$

The $z_{(t)}$ is ($N \times 1$) vector observations at the t-time. Then, λ_k is spatial order from k-th AR, and ϕ_{kl} is a diagonal matrix with diagonal elements as autoregressive (AR) and space-time for each location ($\Phi_{kl}^{(1)}, \ldots, \Phi_{kl}^{(N)}$). Last, $\varepsilon_{(t)}$ is white noise with an average vector of 0 and a matrix of variance-covariance $\sigma^2 I$ (Borovkova, Lopuhaä, & Ruchjana, 2008).

GSTARX model is a GSTAR model that considers the existence of exogenous variables. The variables are thought to affect modeled endogenous variables. GSTARX model with the autoregressive order *(p)*, spatial order *(λ)*, and the order transfer function *(b, r, s)* can be written as follows:

$$z_{(t)} = \sum_{k=1}^{p} \sum_{l=0}^{\lambda_k} \phi_{kl} W^{(l)} z_{(t-k)} + \omega(B)[\delta(B)]^{-1} X_{(t-b)} + \varepsilon_{(t)} \qquad (2)$$

The $X_{(t-b)}$ is $m \times 1$ variable vector X at $t - b$ time, and $\omega(B)[\delta(B)]^{-1}$ is $m \times m$ diagonal matrix of the transfer function parameters, with $\omega(B) = \omega_0 - \omega_1 B - \cdots - \omega_s B^s$ and $\delta(B) = I - \delta_1 B - \cdots - \delta_r B^r$.

Then, SUR is an equation parameter estimation using GLS (Nisak, 2016). SUR model consists of several equations in which the remainder does not correlate between observations in one equation. However, it relates one equation with another. To test the correlation between the equations, the Lagrange Multiplier test statistic is used (Greene, 2012). SUR model with m equations is stated as follows (Nisak, 2016). It shows Z as observations vector, X as diagonal matrix of independent variables, β as parameter model, and ε as residuals vector.

$$Z = X\beta + \varepsilon \qquad (3)$$

Assumptions that must be fullfiled in the SUR model are $E(\varepsilon) = 0$ and $E(\varepsilon\varepsilon') = \sigma_{ij} I_T$ where $E(\varepsilon)$ is the expected value of residuals vector, $E(\varepsilon\varepsilon')$ is the expected value of residuals vector multiplication, σ_{ij} is variance matrix, I_T is identity matrix, and $i, j = 1, 2, \ldots, m$. Variance-covariance matrix stated by Ω is in Equation (4). The matrix Ω sized is $(N \times T) \times (N \times T)$.

$$\Omega = \Sigma \otimes I_T \qquad (4)$$

According to Setiawan, Suhartono, and Prastuti (2016), parameter estimation in GSTARX-SUR model is done by applying the GLS method. It minimizes the number of general squares of $\varepsilon'\Omega - I\varepsilon$. The results from GLS estimators for GSTAR-SUR and GSTARX-SUR models are obtained with Equation (5). The $\hat{\beta}$ is parameter estimators, and Ω is variance-covariance matrix.

$$\hat{\beta} = (X'\Omega^{-1}X)^{-1} X'\Omega^{-1}Z \qquad (5)$$

Next, Kriging interpolation is a mathematical function to estimate values at locations where the data are not available or not sampled based on the sampled points around it using the suitable semivariogram model. Semivariogram describes and models the spatial autocorrelation between data from a variable and functions as a variance measure (Gaetan & Guyon, 2010). Semivariogram is divided into two kinds: theoretical semivariogram and experimental semivariogram. There are various theoretical semivariogram models, such as spherical, exponential, gaussian, and circular models. An experimental semivariogram is based on the spatial correlation value between two variables separated by a certain distance. The experimental semivariogram is formulated in Equation (6) (Cressie, 1993). It shows s_i as sample point location; $Z(s_i)$ as observation value at the location of s_i; h as the distance between two sample points of s_i; $s_i + h$ as pair of sample points within h; and $N(h)$ as the number of data pairs that have a distance of h.

$$\gamma^*(h) = \frac{1}{2N(h)} \sum_{i=1}^{N(h)} [Z(s_i + h) - Z(s_i)]^2 \qquad (6)$$

After the experimental semivariogram values are obtained, parameters can be calculated for theoretical semivariogram calculations. Some parameters used to find values in theoretical semivariograms are sill, nugget, and range (Webster & Oliver, 1992). After obtaining the values of the three parameters, the theoretical semivariogram values are calculated and compared with the experimental semivariogram.

The research uses GSTAR model with SUR approach and tries to add exogenous variables to the model and integrate it with Kriging interpolation techniques in forecasting in unobserved locations. The case study is forecasting coffee borer beetle in Probolinggo Regency. This case study is chosen by considering that coffee is an essential commodity for the national economy and has a high economic value. The problems in the coffee industry are the low productivity of coffee plants and the low quality of coffee beans. The leading cause is the high attack of pests and diseases. One of the main pests that attack coffee plants is the coffee borer beetle. The damage by this pest can reach 4050% of the weight of coffee beans. One of the centers of coffee borer infestation in East Java is Probolinggo Regency.

The coffee borer beetle is a major pest in coffee plantations worldwide (Infante, Pérez, & Vega, 2012). It causes a decrease in productivity and quality of the beans. It is a dark black beetle borer known as *Hypothenemus Hampei* Ferr. It is also a major plant pest of coffee plants because of its rapid development (Baker, Barrera, & Rivas, 1992). It attacks young fruits and causes fruit declines. Meanwhile, the attacks on fairly old fruit cause defective holes and poor quality coffee beans (Damon, 2000; Jaramillo, Borgemeister, & Baker, 2006). It means that fruit borer attacks cause low production and reduce the quality of coffee beans, which results in increased costs for sorting defective seeds. Population dynamics and patterns of infestation by the coffee borer beetle are closely related to climate factors such as rainfall and relative humidity, and

the physiology of coffee plants (Jaramillo *et al.,* 2006).

II. METHODS

The research is quantitative research. The data consist of secondary data and primary data. Secondary data are used to build the model (training data), and primary data validate the model obtained. The training data are 66 observations from January 2014 to June 2019. Meanwhile, the testing data are three observations from July to September 2019. Data are multivariate time series in ten villages in the six biggest coffee producing sub-districts in Probolinggo Regency. Secondary data are obtained from Balai Besar Perbenihan dan Proteksi Tanaman Perkebunan (BBPPTP) in Surabaya and the official website of the National Aeronautics and Space Administration (NASA). Then, the primary data are obtained from direct observation of the research location. The research variable is the percentage of the coffee borer beetle monthly attack area (in percentage) as a predictor variable and monthly rainfall (in millimeters) as an exogenous variable.

There are six steps of the data analysis in the research. First, perform the GSTAR and GSTARX model by identifying the GSTAR and GSTARX model orders and estimating the parameters of the GSTAR-SUR and GSTARX-SUR models. Second, forecast the next three months using the obtained GSTAR-SUR and GSTARX-SUR models. Third, model the GSTAR-SUR Kriging and GSTARX-SUR Kriging by determining nine observed locations and one unobserved location, estimating the parameters of the GSTARX-SUR model at observed locations with SUR and unobserved locations with Kriging interpolation, and establishing GSTARX-SUR Kriging forecasting model in all locations. Fourth, forecast the next three months using the obtained GSTAR-SUR Kriging and GSTARX-SUR Kriging models. Fifth, determine the best forecasting model based on the Root Mean Squared Error (RMSE) and Mean Absolute Percentage Error (MAPE) values. Last, make a forecast map of coffee borer beetle attacks in the Probolinggo Regency.

III. RESULTS AND DISCUSSIONS

The spatial order in the research is limited to $\lambda = 1$, and the order of time *(p)* is determined through the Matrix Partial Cross Correlation Function (MPCCF) and Corrected Akaike's Information Criterion (AICC) schemes. Based on the MPCCF and AICC schemes, it can be concluded that the order is $p = 1$. In addition to lag 1, GSTAR model time order also adds lag 12 by considering that data plots for coffee borer beetle attacks tend to have a 12-month seasonal pattern. Hence, the GSTAR model *(λ, p)* with spatial lag 1 is expressed as GSTAR (1, [1,12]). Furthermore, the order of the transfer function is identified using the Cross-Correlation Function (CCF) plot. The structure of the GSTARX-SUR model can be expressed by GSTARX-SUR (1, [1,12]) (10,0,0).

Estimation of GSTAR model parameters is done by using the OLS method first to generate the residuals. From the residuals of the OLS method, the obtained matrix of variance-covariance will be used in estimating model parameters using the SUR method. In summary, the estimation results of the GSTAR-SUR model parameters are presented in Table 1.

Table 1 The Estimated Results
of the GSTAR-SUR Model Parameters

Location	β-Parameter	β-Estimate	P-Value
Z_1	$\varphi_{10}^{(1)}$	0,452	0,000
	$\varphi_{120}^{(1)}$	0,269	0,032
	$\varphi_{11}^{(1)}$	0,156	0,209
	$\varphi_{121}^{(1)}$	0,120	0,316
Z_2	$\varphi_{10}^{(2)}$	0,282	0,084
	$\varphi_{120}^{(2)}$	0,072	0,616
	$\varphi_{11}^{(2)}$	0,348	0,070
	$\varphi_{121}^{(2)}$	0,298	0,098
Z_3	$\varphi_{10}^{(3)}$	0,434	0,002
	$\varphi_{120}^{(3)}$	0,007	0,962
	$\varphi_{11}^{(3}$	0,191	0,185
	$\varphi_{121}^{(3)}$	0,370	0,018
Z_4	$\varphi_{10}^{(4)}$	0,167	0,221
	$\varphi_{120}^{(4)}$	0,173	0,229
	$\varphi_{11}^{(4)}$	0,491	0,001
	$\varphi_{121}^{(4)}$	0,135	0,350
Z_5	$\varphi_{10}^{(5)}$	0,274	0,057
	$\varphi_{120}^{(5)}$	-0,116	0,517
	$\varphi_{11}^{(5)}$	0,384	0,012
	$\varphi_{121}^{(5)}$	0,429	0,013
Z6	$\varphi_{10}^{(6)}$	0,151	0,111
	$\varphi_{120}^{(6)}$	-0,088	0,366
	$\varphi_{11}^{(6)}$	0,563	0,000
	$\varphi_{121}^{(6)}$	0,384	0,001
Z_7	$\varphi_{10}^{(7)}$	0,602	0,001
	$\varphi_{120}^{(7)}$	0,224	0,131
	$\varphi_{11}^{(7)}$	0,086	0,640
	$\varphi_{121}^{(7)}$	0,068	0,689
Z_8	$\varphi_{10}^{(8)}$	0,508	0,002
	$\varphi_{120}^{(8)}$	0,274	0,119
	$\varphi_{11}^{(8)}$	0,176	0,311
	$\varphi_{121}^{(8)}$	0,026	0,897
Z_9	$\varphi_{10}^{(9)}$	0,692	0,000
	$\varphi_{120}^{(9)}$	-0,138	0,070
	$\varphi_{11}^{(9)}$	0,027	0,807
	$\varphi_{121}^{(9)}$	0,412	0,001
Z_{10}	$\varphi_{10}^{(10)}$	0,610	0,000
	$\varphi_{120}^{(10)}$	-0,269	0,022
	$\varphi_{11}^{(10)}$	0,058	0,657
	$\varphi_{121}^{(10)}$	0,624	0,000

From the estimation results of the GSTAR model parameters, there are locations where the coffee borer beetle attack is affected by the attack one month earlier at that location. Meanwhile, some attacks are affected by the attack in one month and one year earlier at that location, and

some are affected by the one attack in the previous month at the surrounding location. It shows that the coffee borer beetle attack in Probolinggo Regency is influenced not only by the time aspect but also by the spatial aspect.

Similar to the GSTAR, the GSTARX model parameter estimation is performed using the OLS method to generate the remainder. Furthermore, from the remainder of the OLS method, a matrix of variance-covariance will estimate the model parameters using the SUR method. In summary, the estimation results of GSTARX-SUR model parameters are presented in Table 2.

Table 2 The Estimated Results
of The GSTARX-SUR Model Parameters

Location	β-Parameter	β-Estimate	P-Value
Z_1	$\varphi_{10}^{(1)}$	0,389	0,000
	$\varphi_{120}^{(1)}$	0,239	0,065
	$\varphi_{11}^{(1)}$	0,206	0,171
	$\varphi_{121}^{(1)}$	0,134	0,315
	$\omega_{10}^{(1)}$	0,007	0,645
Z_2	$\varphi_{10}^{(2)}$	0,244	0,099
	$\varphi_{120}^{(2)}$	0,033	0,805
	$\varphi_{11}^{(2)}$	0,316	0,068
	$\varphi_{121}^{(2)}$	0,258	0,118
	$\omega_{10}^{(2)}$	0,033	0,029
Z_3	$\varphi_{10}^{(3)}$	0,341	0,011
	$\varphi_{120}^{(3)}$	0,027	0,834
	$\varphi_{11}^{(3)}$	0,214	0,119
	$\varphi_{121}^{(3)}$	0,274	0,048
	$\omega_{10}^{(3)}$	0,033	0,010
Z_4	$\varphi_{10}^{(4)}$	0,162	0,233
	$\varphi_{120}^{(4)}$	0,165	0,245
	$\varphi_{11}^{(4)}$	0,426	0,005
	$\varphi_{121}^{(4)}$	0,064	0,650
	$\omega_{10}^{(4)}$	0,033	0,002
Z_5	$\varphi_{10}^{(5)}$	0,267	0,036
	$\varphi_{120}^{(5)}$	-0,119	0,445
	$\varphi_{11}^{(5)}$	0,314	0,021
	$\varphi_{121}^{(5)}$	0,353	0,019
	$\omega_{10}^{(5)}$	0,035	0,001
Z_6	$\varphi_{10}^{(6)}$	0,139	0,127
	$\varphi_{120}^{(6)}$	-0,088	0,359
	$\varphi_{11}^{(6)}$	0,485	0,000
	$\varphi_{121}^{(6)}$	0,298	0,008
	$\omega_{10}^{(6)}$	0,039	0,004
Z_7	$\varphi_{10}^{(7)}$	0,622	0,002
	$\varphi_{120}^{(7)}$	0,452	0,012
	$\varphi_{11}^{(7)}$	-0,004	0,985
	$\varphi_{121}^{(7)}$	-0,277	0,178
	$\omega_{10}^{(7)}$	0,043	0,003
Z_8	$\varphi_{10}^{(8)}$	0,497	0,005
	$\varphi_{120}^{(8)}$	0,294	0,110
	$\varphi_{11}^{(8)}$	0,098	0,589
	$\varphi_{121}^{(8)}$	-0,071	0,736
	$\omega_{10}^{(8)}$	0,035	0,005
Z_9	$\varphi_{10}^{(9)}$	0,658	0,000
	$\varphi_{120}^{(9)}$	-0,163	0,046
	$\varphi_{11}^{(9)}$	-0,041	0,726
	$\varphi_{121}^{(9)}$	0,331	0,004
	$\omega_{10}^{(9)}$	0,045	0,005
Z_{10}	$\varphi_{10}^{(10)}$	0,620	0,000
	$\varphi_{120}^{(10)}$	-0,303	0,011
	$\varphi_{11}^{(10)}$	-0,033	0,812
	$\varphi_{121}^{(10)}$	0,598	0,000
	$\omega_{10}^{(10)}$	0,033	0,052

From the estimation results of the GSTARX-SUR model parameters in Table 2, the results of the significance testing of the GSTARX-SUR model parameters between locations also vary considerably. Based on the results of the parametric test of ω_{10}, the rainfall in ten months earlier significantly influences the coffee borer beetle attack in nine locations. It shows that the coffee borer beetle attack in Probolinggo Regency is affected by the time and spatial aspects and rainfall from the previous ten months. Based on the estimation results of the model parameters in Table 1 and Table 2, GSTAR-SUR and GSTARX-SUR models can predict the coffee borer beetle attacks in ten villages in six sub-districts in the Probolinggo Regency. Next, using the GSTAR-SUR and GSTARX-SUR models, the forecasting of coffee borer beetle attack can be done. Figure 1 presents a plot of prediction data and actual data in Segaran Village, Tiris Sub-district.

In the research, from ten villages in Probolinggo Regency, there are nine villages with completely available data. However, the data of Watupanjang Village, Krucil Sub-district village are not available. This village is chosen as an unobserved location. The location of coffee plantations in the area is difficult to reach due to difficult road access. GSTARX-SUR model in nine locations is observed using stages as well as GSTARX-SUR model in ten locations. The forecasting stage is done after the Kriging interpolation process to get the estimated value of the GSTARX-SUR Kriging model parameters in an unobserved location.

Estimating model parameters at unobserved locations is done by Kriging interpolation using the estimated values of model parameters at nine observed locations. Comparison of the estimated results of the GSTAR Kriging and GSTARX Kriging model parameters with the ordinary GSTAR and GSTARX models in Watupanjang Village is presented in Table 3 and Table 4.

Table 3 The Comparison of Estimated Parameter
Values between the GSTAR-SUR
and GSTAR-SUR Kriging Models

Parameter	GSTAR Kriging	GSTAR
$\varphi_{10}^{(7)}$	0,617	0,602
$\varphi_{120}^{(7)}$	0,228	0,224
$\varphi_{11}^{(7)}$	0,053	0,086
$\varphi_{121}^{(7)}$	0,084	0,068

Table 4 Comparison of Estimated Parameter
Values between the GSTARX-SUR
and GSTARX-SUR Kriging Models

Parameter	GSTARX Kriging	GSTARX
$\varphi_{10}^{(7)}$	0,553	0,622
$\varphi_{120}^{(7)}$	0,234	0,452
$\varphi_{11}^{(7)}$	0,046	-0,004
$\varphi_{121}^{(7)}$	-0,021	-0,277
$\omega_{10}^{(7)}$	0,037	0,043

In Table 3, the results of the estimated parameters of the GSTAR-SUR Kriging model do not differ greatly from the ordinary GSTAR-SUR model. Meanwhile, in Table 4, the estimation results of the GSTARX-SUR Kriging model parameters for the three parameters are relatively not much different, and the other two parameters are slightly different. However, the difference is not so big. The results obtained are similar to Abdullah *et al.* (2018). The estimation results

of the model parameters in both the observed locations and one unobserved location are used to form the GSTAR-SUR Kriging and GSTARX-SUR Kriging models to predict the coffee borer beetle attacks in ten locations.

The goodness of the GSTAR-SUR Kriging and GSTARX-SUR Kriging can be seen by looking at the forecasting of MAPE value. It can also compare RMSE values between the GSTAR-SUR Kriging and GSTAR-SUR and GSTARX-SUR Kriging and GSTARX-SUR. Forecasting models are reliable if they have a MAPE forecasting value of less than 10%. The best model selection criteria are performed using MAPE and RMSE values. The best model is the model with smaller MAPE and RMSE values. Table 5 presents the MAPE and RMSE values of the GSTAR-SUR Kriging and GSTARX-SUR Kriging models.

Table 5 MAPE and RMSE values between GSTAR-SUR
Kriging and GSTARX-SUR Kriging Models

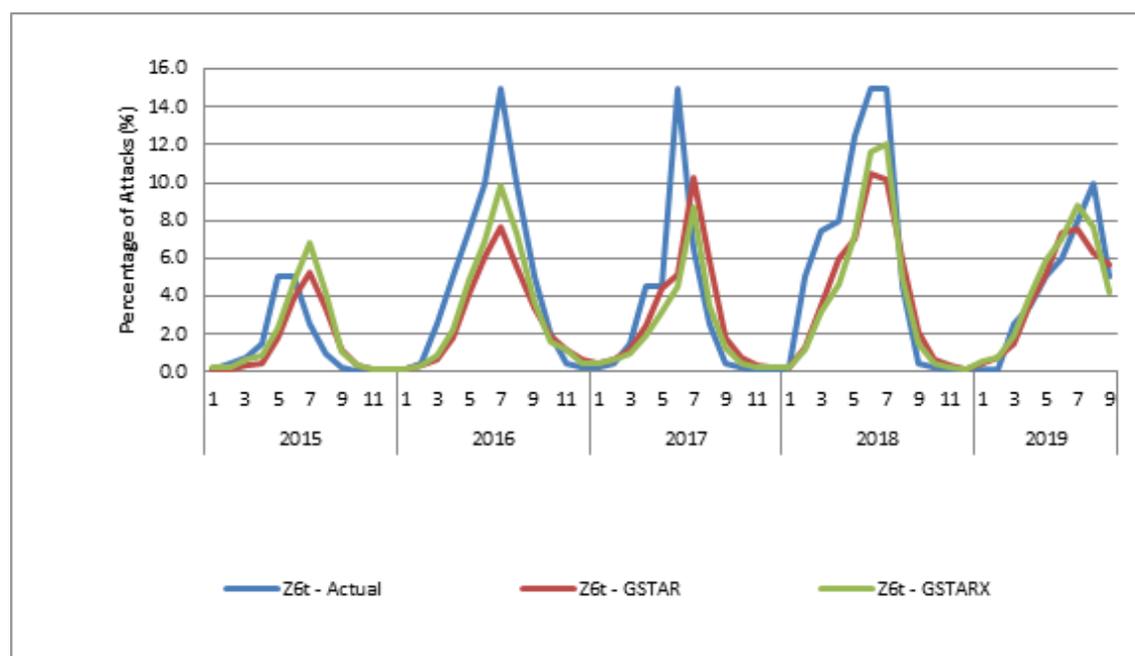| Location	GSTAR Kriging		GSTARX Kriging	
	MAPE	RMSE	MAPE	RMSE
Z1	7,20%	0,0444	**6,75%**	**0,0430**
Z2	11,85%	0,0741	**9,96%**	**0,0716**
Z3	9,88%	**0,0630**	**8,81%**	0,0650
Z4	5,92%	0,0363	**4,97%**	**0,0320**
Z5	5,68%	0,0353	**4,82%**	**0,0298**
Z6	5,20%	0,0380	**4,32%**	**0,0277**
Z7	7,35%	0,0486	**4,98%**	**0,0320**
Z8	7,58%	0,0545	**6,97%**	**0,0536**
Z9	**2,83%**	**0,0197**	4,37%	0,0298
Z10	**2,79%**	**0,0206**	5,80%	0,0382
Overall	6,63%	0,0434	**6,18%**	**0,0423**

Figure 1 Actual Data and Prediction Data Plots using GSTAR-SUR
and GSTARX-SUR Models in Segaran Village, Tiris Sub-district

In Table 5, both GSTAR-SUR Kriging and GSTARX-SUR Kriging have MAPE values of less than 10%, so both of them are appropriate to predict coffee borer beetle attacks in Probolinggo Regency. When those are compared, the accuracy of the GSTARX-SUR Kriging is better than the GSTAR-SUR Kriging. It is indicated by the overall MAPE and RMSE values of the GSTARX-SUR Kriging, which are smaller than the GSTAR-SUR Kriging.

The researchers then map the forecasting of coffee borer beetle attacks in eight coffee-producing districts in Probolinggo Regency. It is the forecasting result of interpolation of the coffee borer beetle attack in ten locations in six coffee-producing districts in the Probolinggo Regency

and the results of the GSTAR-SUR Kriging and GSTARX-SUR Kriging models. From each forecasting model, two forecasting maps are formed, namely the forecasting map of coffee borer beetle attack for July 2019 and August 2019. Meanwhile, forecasting results for September 2019 are not mapped because the attacks are relatively low and evenly distributed in all districts. The map of forecasting coffee borer attack in July and August 2019 from GSTAR-SUR Kriging is presented in Figures 2 and 3.

From the forecasting map of the coffee borer beetle attacks resulting from the GSTAR Kriging model, coffee borer beetle attacks in July are predicted to be quite high in the eastern region of Probolinggo Regency, especially in
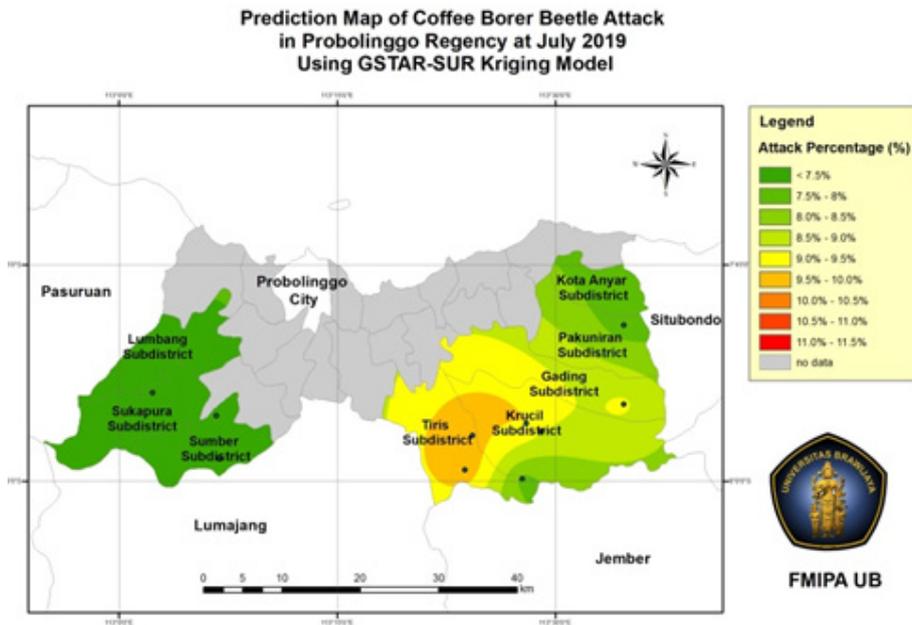

Figure 2 Prediction Map of Coffee Borer Beetle Attack in Probolinggo Regency
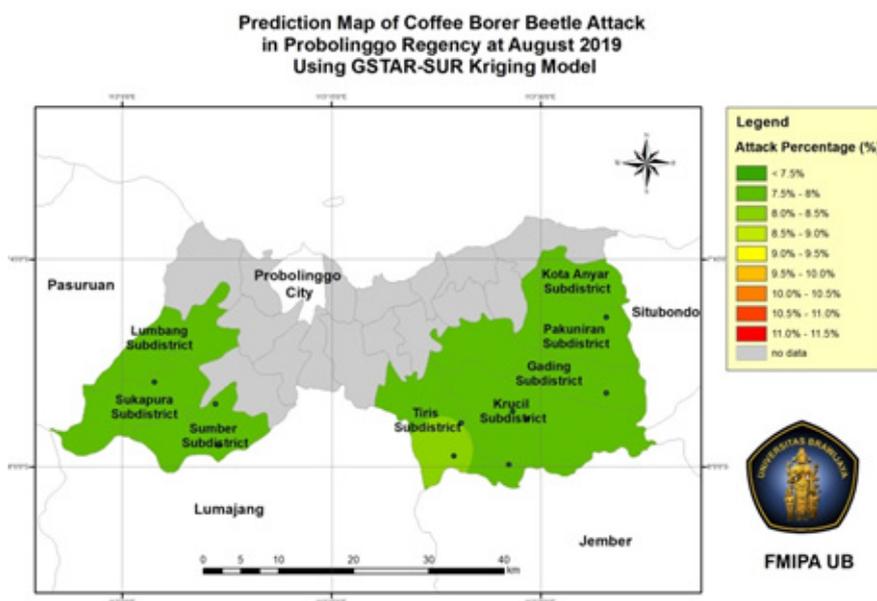in July 2019 using GSTAR-SUR Kriging Model

Figure 3 Prediction Map of Coffee Borer Beetle Attack in Probolinggo Regency
in August 2019 using GSTAR-SUR Kriging Model

the Tiris area. The coffee borer beetle attack in the eastern area of Probolinggo Regency is relatively higher than the western area. It considers that the type of coffee planted in the eastern area is generally Robusta coffee, which is more susceptible to the coffee borer beetle attacks. The coffee borer beetle attacks are predicted to begin to decline in August in almost all regions. The coffee borer beetle attacks in September are predicted to decrease, and the intensity of the attack is no more than 7%. The map of forecasting coffee borer beetle attack for July and August 2019 using GSTARX-SUR Kriging model is presented in Figures 4 and 5.

In Figures 4 and 5, the coffee borer beetle attacks in July are predicted to be quite high in the eastern region of Probolinggo Regency, especially in the Tiris, Krucil, and Gading areas. The coffee borer beetle attacks in the eastern region are relatively higher than in the western region. The coffee borer beetle attacks are predicted to start to decline in August in almost all regions. Only in the southern part of Tiris, which has relatively high attack intensity. The coffee borer beetle attacks in September are predicted to decrease, and the intensity of the attack is no more than 7%. Comparing the forecasting map, the GSTARX-SUR Kriging provides a slightly higher forecasting value than the GSTAR-SUR Kriging.
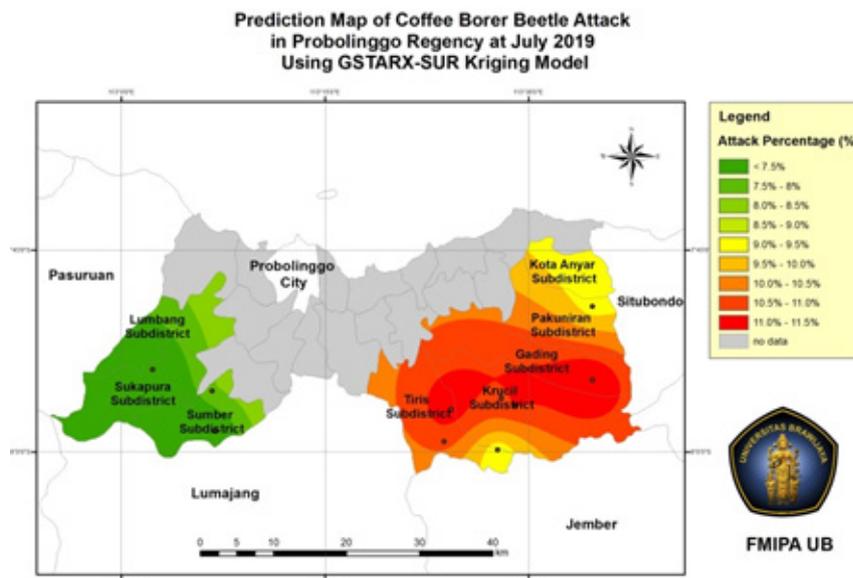

Figure 4 Prediction Map of Coffee Borer Beetle Attack in Probolinggo Regency in July 2019 using GSTARX-SUR Kriging Model
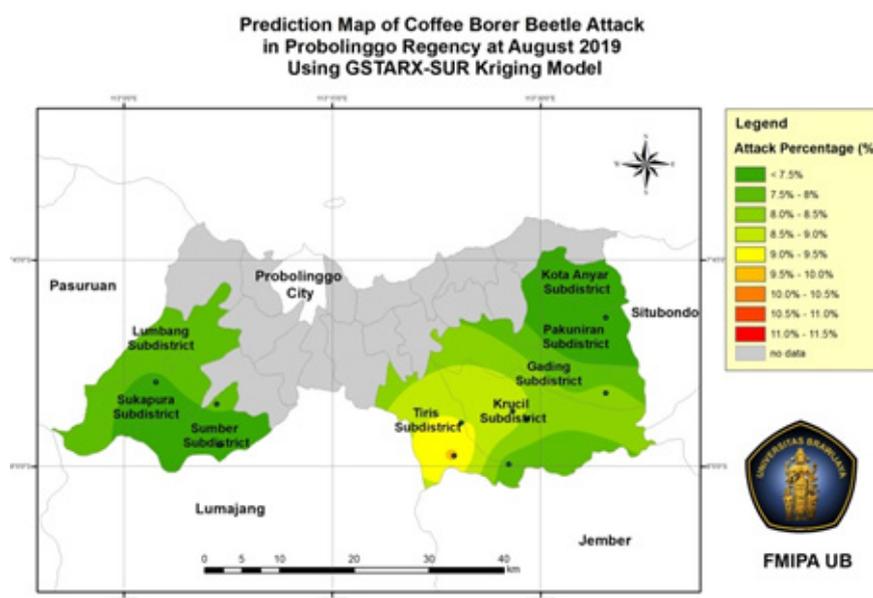

Figure 5 Prediction Map of Coffee Borer Beetle Attack in Probolinggo Regency in August 2019 using GSTARX-SUR Kriging Model

Based on the forecasting map of coffee borer beetle attacks, it is recommended to the Probolinggo Regency plantation office together with farmers to carry out Integrated Pest Control (IPM) during June. It is to prevent the high intensity of the coffee borer beetle attacks, which are predicted to occur during July in the Tiris, Krucil, and Gading areas. It also needs to be done considering that the distribution pattern of coffee borer beetle attacks tends to be in a seasonal pattern. Thus, the successful pest control this year will affect the coffee borer beetle attacks in the next year's harvest. Another effort that can be done is by cleaning the remaining attacked coffee fruit (not harvested) and not leaving it to break the life cycle of the coffee borer beetle attacks. The long dry season during 2019 also has the potential to slow down the flowering phase. It is estimated that the harvest period in 2020 will be delayed so that the peak of the attack next year slightly changes. If the usual peak of the coffee borer beetle attacks occurs in June and July, in 2020, it will shift slightly to July and August 2020.

IV. CONCLUSIONS

GSTAR-SUR Kriging and GSTARX-SUR Kriging models can predict coffee borer beetle attacks in unobserved locations with high accuracy. It is indicated by MAPE values of less than 10%. The addition of exogenous variables (rainfall) into the model is proven to improve the accuracy of the model. The RMSE value of the GSTARX-SUR Kriging model is smaller than the GSTAR-SUR Kriging model. The structure of the model produced from the research, GSTARX-SUR (1,[1,12])(10,0,0), can be used as a reference in modeling coffee borer attacks in other regencies. Map of forecasting coffee borer beetle attack shows that the spread is spatial clustering with the attack center located in the eastern region of Probolinggo Regency.

The GSTAR-SUR Kriging and GSTARX-SUR Kriging models have limitations. The forecasting in an unobserved location still requires observational data at that location. The data can be observational data at one and twelve months before. Therefore, for further research, modifications can be made in the flow of interpolation. Future researchers can conduct Kriging interpolation first to predict time series data in unobserved locations. After completing data in all locations, GSTAR-SUR or GSTARX-SUR models can be formed as usual. In addition, they can also combine the GSTARX-SUR model with other interpolation techniques besides Kriging, such as Inverse Distance Weighted (IDW) or Spline.

REFERENCES

Abdullah, A. S, Matoha, S., Lubis, D. A., Falah, A. N., Jaya, I. G. N. M., Hermawan, E., & Ruchjana, B. N. (2018). Implementation of Geralized Space Time Autoregressive (GSTAR)-Kriging model for predicting rainfall data at unobserved locations in West Java. *Applied Mathematics & Information Sciences, 12*(3), 607-615. https://doi.org/10.18576/amis/120316

Andayani, N., Sumertajaya, I. M., Ruchjana, B. N., & Aidi, M. N. (2017). Development of space time model with exogenous variable by using transfer function model approach on the rice price data. *Applied Mathematical Sciences, 11*(36), 1779-1792. https://doi.org/10.12988/ams.2017.74150

Astuti, D., Ruchjana, B. N., & Soemartini. (2017). Generalized Space Time Autoregressive with exogenous variable model and its application. *Journal of Physics: Conference Series, 893*, 1-9. https://doi.org/10.1088/1742-6596/893/1/012038

Baker, P. S., Barrera, J. F., & Rivas, A. (1992). Life-history studies of the coffee berry borer (Hypothenemus Hampei, Scolytidae) on coffee trees in Southern Mexico. *Journal of Applied Ecology, 29*(3), 656-662. https://doi.org/10.2307/2404473

Borovkova, S., Lopuhaä, H. P., & Ruchjana, B. N. (2008). Consistency and asymptotic normality of least squares estimators in generalized STAR models. *Statistica Neerlandica, 62*(4), 482-508. https://doi.org/10.1111/j.1467-9574.2008.00391.x

Cressie, N. A. C. (1993). *Statistics for spatial data* (Revised edition). Hoboken, NJ: John Wiley & Sons.

Damon, A. (2000). A review of the biology and control of the coffee berry borer, Hypothenemus Hampei (Coleoptera: Scolytidae). *Bulletin of Entomological Research, 90(6)*, 453-465. https://doi.org/10.1017/S0007485300000584

Gaetan, C., & Guyon, X. (2010). Statistics for spatial models. In C. Gaetan & X. Guyon (Ed.), *Spatial statistics and modeling* (pp. 149-248). Springer.

Greene, W. H. (2012). *Econometric analysis* (7th ed). Pearson.

Infante, F., Pérez, J., & Vega, F. E. (2012). Redirect research to control coffee pest. *Nature, 489*(7417), 502-502. https://doi.org/10.1038/489502a

Iriany, A., Suhariningsih, Ruchjana, B. N., & Setiawan. (2013). Prediction of precipitation data at Batu Town using the GSTAR(1,p)-SUR model. *Journal of Basic and Applied Scientific Research, 3*(6), 860-865.

Jaramillo, J., Borgemeister, C., & Baker, P. (2006). Coffee berry borer Hypothenemus Hampei (Coleoptera: Curculionidae): Searching for sustainable control strategies. *Bulletin of Entomological Research, 96*(3), 223-233. https://doi.org/10.1079/BER2006434

Nisak, S. C. (2016). Seemingly unrelated regression approach for GSTARIMA model to forecast rain fall data in Malang southern region districts. *CAUCHY: Jurnal Matematika Murni dan Aplikasi, 4*(2), 57-64. https://doi.org/10.18860/ca.v4i2.3488

Ruchjana, B. N., Borovkova, S. A., & Lopuhaa, H. P. (2012). Least squares estimation of Generalized Space Time AutoRegressive (GSTAR) model and its properties. *AIP Conference Proceedings, 1450*, 61-64. https://doi.org/10.1063/1.4724118

Setiawan, Suhartono, & Prastuti, M. (2016). S-GSTAR-SUR model for seasonal spatio temporal data forecasting. *Malaysian Journal of Mathematical Sciences, 10*(S), 53-65.

Suhartono, Wahyuningrum, S. R., Setiawan, & Akbar, M. S. (2016). GSTARX-GLS model for spatio-temporal data forecasting. *Malaysian Journal of Mathematical Sciences, 10*(S), 91-103.

Webster, R., & Oliver, M. A. (1992). Sample adequately to estimate variograms of soil properties. *Journal of Soil Science, 43*(1), 177-192. https://doi.org/10.1111/j.1365-2389.1992.tb00128.x