Cross-Covariance Weight of GSTAR-SUR Model for Rainfall Forecasting in Agricultural Areas

Agus Dwi Sulistyono[1], Hartawati[2], Ni Wayan Suryawardhani[3] , Atiek Iriany[3] , Aniek Iriany[2]

[1] Faculty of Fisheries and Marine Science, Brawijaya University, Indonesia
[2] Department of Agrotechnology, Faculty of Agriculture and Animal Husbandry, University of Muhammadiyah Malang, Indonesia
[3] Department of Statistics Faculty of Mathematics and Natural Sciences, Brawijaya University, Indonesia

Email: agus.dwi.sulistyono@gmail.com

ABSTRACT

The use of location weights on the formation of the spatio-temporal model contributes to the accuracy of the model formed. The location weights that are often used include uniform location weight, inverse distance, and cross-correlation normalization. The weight of the location considers the proximity between locations. For data that has a high level of variability, the use of the location weights mentioned above is less relevant. This research was conducted with the aim of obtaining a weighting method that is more suitable for data with high variability. This research was conducted using secondary data derived from 10 daily rainfall data obtained from BMKG Karangploso. The data period used was January 2008 to December 2018. The points of the rain posts studied included the rain post of the Blimbing, Karangploso, Singosari, Dau, and Wagir regions. Based on the results of the research forecasting model obtained is the GSTAR ((1), 1,2,3,12,36) -SUR model. The cross-covariance model produces a better level of accuracy in terms of lower RMSE values and higher R^2 values, especially for Karangploso, Dau, and Wagir areas.

Keywords: cross-covariance; GSTAR Model; rainfall; spatio-temporal

INTRODUCTION

There are several spatio-temporal models that have been developed. For the first time, Space-Time Autoregressive (STAR) model was introduced by Pfeifer & Deutsch [1], [2]. The Space-Time Autoregressive (STAR) model had the assumption that the variance between locations is the same/homogeneous. However, in fact, it often gets heterogeneity between observation sites. Thus the STAR model is not suitable for data that has heterogeneous location characteristics. This is the weakness of the STAR model and this weakness can be handled by the Generalized Space-Time Autoregressive (GSTAR) and GSTAR-OLS models developed by Borovkova, Lopuha, & Ruchjana [3] and Ruchjana [4], [5]. The GSTAR model developed is used for data that meets stationary assumptions. The latest development of this spatio-temporal model is the GSTAR-SUR model developed by Iriany [6] to overcome data that is not stationary and has a seasonal pattern. Furthermore, the use of the GSTAR-SUR-NN hybrid model was also developed for data that has a nonlinear pattern [7]–[9]

The use of location weights in the formation of spatiotemporal models also

contributes to the level of accuracy of the model formed. There are some types of location weights that are used to build models, there are uniform location weight, inverse distance, and cross-correlation normalization [10], [11]. The weight of the location considers the proximity between locations. For data that has a high level of variability, the use of the location weights mentioned above is less relevant. Therefore, we need location weights that consider the various aspects of observational data. One of the location weights that have been developed is the weight of the variance ratio which has proven to have a better level of accuracy [12]. The weight of other locations developed is the weight of cross-covariance. The use of the weight of cross-covariance has been researched and applied in the research of Apanosovich and Genton [13] to predict pollution in California and the research of Efromovich & Smirnova [14] for fMRI imaging processes with a wavelet approach. This research was conducted to determine the accuracy of the GSTAR model that was built using the weight of cross-covariance and compare the level of accuracy with the GSTAR model that was built with the weight of cross-correlation.

METHODS

The data period used is January 2008 to December 2018, where data for conducting a training (in-sample) is data from January 2008 to December 2017. While data from January 2018 to December 2018 is used as testing data (out-of-sample). The first step taken is testing the stationary data on rainfall. The stationary test on the average is done using the Augmented Dickey-Fuller test. While the stationarity test for the variance was carried out by the Box-Cox test. The next step is to identify the real MACF and MPACF lags to determine the order that will be used as an estimate of the GSTAR model. Next, the cross-correlation normalization weighting is calculated with Equation 1 [11]:

$$w_{ij} = \frac{r_{ij}(1)}{\sum_{k \neq i} |r_{ik}(1)|}$$

(1)

and the normalization weight of cross-covariance is calculated with Equation 2 [13], [14] :

$$\sum_{t=k+1}^{n} \left[z_i(t) - \overline{z}_i \right]\left[z_j(t-k) - \overline{z}_j \right] = r_{ij}(k)\sqrt{\left(\sum_{t=1}^{n} \left[z_i(t) - \overline{z}_i \right]^2 \right)\left(\sum_{t=1}^{n} \left[z_j(t-k) - \overline{z}_j \right]^2 \right)}$$

(2)

The next process is GSTAR-OLS analysis to get the residual value with Equation 3

$$Z(t) = \mu(t) + \left[\Phi_{01} + \Phi_{11}W \right]Z(t-1) + \varepsilon(t)$$

(3)

Next, calculate the var $(\varepsilon) = \Omega$ matrix with the equation

$$\Sigma = \begin{bmatrix} \sigma_{11} & \sigma_{12} & \cdots & \sigma_{1m} \\ \sigma_{21} & \sigma_{22} & \cdots & \sigma_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{m1} & \sigma_{m2} & \cdots & \sigma_{mm} \end{bmatrix}$$

(4)

The next step is estimating the GSTAR (1, p) -SUR parameter using the formula $\beta = (X'\Omega^{-1}X)^{-1}X'\Omega^{-1}y$. The best model is chosen based on RMSE and R^2 prediction values. The research data analysis process was carried out using R and SAS software.

RESULTS AND DISCUSSION

This research was carried out by taking daily rainfall data obtained from the rain heading point for the Blimbing, Karangploso, Singosari, Dau, and Wagir regions. The following is a description of the statistics of rainfall data in the five locations presented in Table 1:

Table 1. Description of Rainfall Data Statistics in Five Research Locations

Location	N	Mean (mm)	Standard Deviation (mm)	Minimum (mm)	Maximum (mm)
Blimbing	360	5.682	6.909	0	33.5
Singosari	360	3.93	5.575	0	41.75
Karangploso	360	4.302	5.71	0	25.36
Dau	360	4.564	5.825	0	36.38
Wagir	360	7.08	8.187	0	43.63

Based on Table 1 above, it is descriptively shown that the average rainfall in Wagir District is the highest and Singosari District has the lowest average rainfall. In all study locations, the standard deviation value was greater than the average, indicating a high level of rainfall variation in all study locations. In addition, the heterogeneity of the observation location can be measured by calculating the Gini Index. The higher the index value, the more heterogeneous the location will be. This index calculation for the five locations in this study is:

$$G_n = 1 + \frac{1}{n} - \frac{2}{n^2 \bar{y_i}} \sum_{i=1}^{N} y_i$$

$$= 1 + \frac{1}{360} - \frac{2}{360^2 \, 5.111} \, 9200.68$$

$$= 0.975$$

Based on the results of the Gini index calculation, the Gini index value is 0.975, close to 1. From the Gini Index calculation, it is shown that heterogeneous locations so that modeling using the GSTAR-SUR model can be done.

Stationary testing of variance was carried out using a Box-Cox plot. The stationarity of variance is said to be fulfilled if the Box-Cox plot results in a value of $\lambda = 1$. However, if the value of $\lambda \neq 1$, then the data transformation process is carried out. The following are the results of stationary testing of the variance in rainfall data for each location:

Table 2. Stationary Testing of the Variance of Rainfall in Each Location

Location	λ	Transformation I	
		Transformation	λ
Blimbing	-0.27	$Z^{-0.27}$	1.0
Singosari	-0.68	$Z^{-0.68}$	1.0
Karangploso	-0.50	$Z^{-0.5}$	1.0
Dau	-0.50	$Z^{-0.5}$	1.0
Wagir	-0.19	$Z^{-0.19}$	1.0

Based on Table 4.3 the initial λ values for all study locations have not been worth 1. This shows that the rainfall data in each location is not yet stationary in variety so that Box-Cox transformation is needed. The Box-Cox I transformation results show the value

of $\lambda = 1$, which means that the data has been stationary to the variety and the transformation is stopped.

In addition to stationary variety, stationary testing is also carried out on the average. Stationary to average testing was carried out using the Augmented Dickey-Fuller (ADF) test. Stationarity on the average is said to be fulfilled if the results of the ADF test obtained the p-value of less than 0.05. If the ADF test results obtained the p-value of more than 0.05, it is necessary to do a differencing process. Following are the results of the ADF test:

Table 3. Stationary Testing on the Average Rainfall of Each Location

Location	t-Statistics	p-value
Blimbing	-5.974	0.000
Singosari	-5.556	0.000
Karangploso	-7.831	0.000
Dau	-7.215	0.000
Wagir	-7.425	0.000

Based on the results of the stationary test on the average using the ADF test in Table 3, at each location p-value was less than 0.05. From this test, it is shown that the stationary data of rainfall on the average has been fulfilled.

The GSTAR model identification process is done by looking at the Matrix Partial Autocorrelation Function (MPACF) scheme.

Table 4. Matrix Partial Autocorrelation Scheme (MPACF)

Variable/Lag	1	2	3	4	5	6	7	8	9	10	11	12
	Schematic Representation of Partial Cross Correlations											
BLIMBING	.+++.	.+...	.....	.....	.....	.....	...−.	.....	.....	....−	.....	.....
SINGOSARI	++...	.+...	.+...	.....	.+...	−....	.+...	−....	.....	.....	.....	.....
KARANGPLOSO	+..+.	+...+	...+.	.....	..+..	...−.	.+...	.....	.....	.....	.....	..+..
DAU	..++.	+...+	...−.	.....	.....	.....	.....	.....	...−.	.−...	.....	.....
WAGIR	+..++	+....	.....	.....	...+.	.....	.....	.....	.....	.....	.....	.....

+ is > 2*std error, - is < -2*std error, . is between

Based on the MPACF matrix scheme in Table 4 it can be seen that there is a real MPACF lag in lag 1 to lag 3. Then in lag 4, there is no significant partial autocorrelation. Then in 5 lags and so on there are some significant partial autocorrelations. Based on the MPACF scheme, it is shown that significant partial autocorrelation is truncated at lag 4. So, the determination of the VAR order (p) is done by looking at the smallest AIC value for real lag. The following is the AIC value in lag 1 to lag 3:

Table 5. AIC values for GSTAR Order Determination

Order	AIC Value
1	14.67906
2	14.52868
3	14.52445

Based on the AIC value in Table 5 it can be seen that the lowest AIC value is obtained in the 3rd order. Thus, the GSTAR model used has a 3rd order. In addition to determining the order with the AIC value, identification of the GSTAR model is also carried out by univariate ACF and PACF plots at each location. Based on the ACF plot it is shown that rainfall data at each location is indicated by seasonal patterns. This can be seen in the ACF

plot which has a repetitive pattern at a certain time lag. Based on the PACF plot in Appendix 8 shows that at some time lag there is a PACF that crosses the 5% boundary line. When combined in 5 locations, it was found that the five locations had PACF that passed 12 and 36 of time lags. Therefore, the results of the identification of seasonal patterns indicated that the appropriate model was GSTAR ((1), 1,2,3,12,36)

This study uses five locations with $n_i^{(1)}$ or the number of locations adjacent to the *i*-th location is 4 locations so that the cross-correlation normalization matrix is as follows:

$$W_{ij} = \begin{bmatrix} 0 & 0.2353 & 0.2294 & 0.2688 & 0.2665 \\ 0.2664 & 0 & 0.2143 & 0.2675 & 0.2518 \\ 0.2734 & 0.2100 & 0 & 0.2765 & 0.2401 \\ 0.2950 & 0.2174 & 0.2327 & 0 & 0.2549 \\ 0.2694 & 0.2235 & 0.2420 & 0.2650 & 0 \end{bmatrix}$$

While the magnitude of cross-covariance normalization weighting based on calculations is as follows:

$$W_{ij} = \begin{bmatrix} 0 & 0.2060 & 0.2056 & 0.2458 & 0.3426 \\ 0.2754 & 0 & 0.1831 & 0.2331 & 0.3084 \\ 0.2847 & 0.1764 & 0 & 0.2427 & 0.2962 \\ 0.3058 & 0.1818 & 0.1993 & 0 & 0.3130 \\ 0.3085 & 0.2065 & 0.2290 & 0.2559 & 0 \end{bmatrix}$$

The results of the estimation of the parameters of the GSTAR ((1), 1,2,3,12,36) -SUR model with the weight of the cross correlation normalization location for Blimbing District are as follows:

$\hat{Z}_{1t} = 0.119\ Z_{1(t-1)} + 0.009\ Z_{2(t-1)} - 0.009\ Z_{3(t-1)} + 0.022\ Z_{4(t-1)} + 0.017\ Z_{5(t-1)} + 0.205\ Z_{1(t-2)} + 0.087\ Z_{2(t-2)} + 0.084\ Z_{3(t-2)} + 0.009\ Z_{4(t-2)} + 0.075\ Z_{5(t-2)} - 0.086\ Z_{1(t-3)} - 0.074\ Z_{2(t-3)} + 0.037\ Z_{3(t-3)} + 0.023\ Z_{4(t-3)} + 0.005\ Z_{5(t-3)} + 0.059\ Z_{1(t-12)} - 0.056\ Z_{2(t-12)} - 0.01\ Z_{3(t-12)} + 0.019\ Z_{4(t-12)} - 0.025\ Z_{5(t-12)} + 0.246\ Z_{1(t-36)} + 0.052\ Z_{2(t-36)} - 0.093\ Z_{3(t-36)} + 0.034\ Z_{4(t-36)} + 0.069\ Z_{5(t-36)}$

While the result of the parameter estimation of the GSTAR ((1), 1,2,3,12,36) -SUR model with cross-covariance normalization location weights is as follows :

$\hat{Z}_{1t} = 0.116\ Z_{1(t-1)} + 0.003\ Z_{2(t-1)} - 0.006\ Z_{3(t-1)} + 0.022\ Z_{4(t-1)} + 0.015\ Z_{5(t-1)} + 0.212\ Z_{1(t-2)} + 0.083\ Z_{2(t-2)} + 0.081\ Z_{3(t-2)} + 0.018\ Z_{4(t-2)} + 0.072\ Z_{5(t-2)} - 0.092\ Z_{1(t-3)} - 0.066\ Z_{2(t-3)} + 0.031\ Z_{3(t-3)} + 0.025\ Z_{4(t-3)} + 0\ Z_{5(t-3)} + 0.06\ Z_{1(t-12)} - 0.048\ Z_{2(t-12)} - 0.016\ Z_{3(t-12)} + 0.016\ Z_{4(t-12)} - 0.023\ Z_{5(t-12)} + 0.256\ Z_{1(t-36)} + 0.045\ Z_{2(t-36)} - 0.085\ Z_{3(t-36)} + 0.025\ Z_{4(t-36)} + 0.068\ Z_{5(t-36)}$

The result of parameter estimation of GSTAR ((1), 1,2,3,12,36) -SUR model with the weight of cross-correlation normalization location for Singosari Subdistrict is as follows:

$\hat{Z}_{2t} = 0.024\ Z_{1(t-1)} + 0.171\ Z_{2(t-1)} - 0.009\ Z_{3(t-1)} + 0.02\ Z_{4(t-1)} + 0.012\ Z_{5(t-1)} + 0.017\ Z_{1(t-2)} - 0.255\ Z_{2(t-2)} + 0.083\ Z_{3(t-2)} + 0.009\ Z_{4(t-2)} + 0.052\ Z_{5(t-2)} + 0.055\ Z_{1(t-3)} + 0.649\ Z_{2(t-3)} + 0.036\ Z_{3(t-3)} + 0.021\ Z_{4(t-3)} + 0.004\ Z_{5(t-3)} - 0.017\ Z_{1(t-12)} + 0.335\ Z_{2(t-12)} - 0.01\ Z_{3(t-12)} + 0.018\ Z_{4(t-12)} - 0.018\ Z_{5(t-12)} + 0.031\ Z_{1(t-36)} - 0.084\ Z_{2(t-36)} - 0.092\ Z_{3(t-36)} + 0.031\ Z_{4(t-36)} + 0.048\ Z_{5(t-36)}$

While the result of the parameter estimation of the GSTAR ((1), 1,2,3,12,36) -SUR model with cross-covariance normalization location weights is as follows :

$\hat{Z}_{2t} = 0.024\ Z_{1(t-1)} + 0.206\ Z_{2(t-1)} - 0.006\ Z_{3(t-1)} + 0.021\ Z_{4(t-1)} + 0.01\ Z_{5(t-1)} + 0.013\ Z_{1(t-2)} - 0.281\ Z_{2(t-2)} + 0.081\ Z_{3(t-2)} + 0.016\ Z_{4(t-2)} + 0.05\ Z_{5(t-2)} + 0.055\ Z_{1(t-3)} + 0.639\ Z_{2(t-3)} + 0.03\ Z_{3(t-3)} +$

$0.023\ Z_{4(t-3)} + 0\ Z_{5(t-3)} - 0.017\ Z_{1(t-12)} + 0.301\ Z_2(t\text{-}12) - 0.015\ Z_{3(t-12)} + 0.015\ Z_{4(t-12)} - 0.016\ Z_{5(t-12)} + 0.025\ Z_{1(t-3)}) - 0.04\ Z_{2(t-36)} - 0.084\ Z_{3(t-36)} + 0.023\ Z_{4(t-36)} + 0.047\ Z_{5(t-36)}$

The result of parameter estimation of GSTAR ((1), 1,2,3,12,36) -SUR model with the weight of cross-correlation normalization location for Karangploso Subdistrict is as follows:

$\hat{Z}_{3t} = 0.033\ Z_{1(t-1)} + 0.008\ Z_{2(t-1)} + 0.411\ Z_{3(t-1)} + 0.023\ Z_{4(t-1)} + 0.016\ Z_{5(t-1)} + 0.023\ Z_{1(t-2)} + 0.081\ Z_{2(t-2)} - 0.318\ Z_{3(t-2)} + 0.01\ Z_{4(t-2)} + 0.072\ Z_{5(t-2)} + 0.076\ Z_{1(t-3)} - 0.069\ Z_{2(t-3)} - 0.056\ Z_{3(t-3)} + 0.024\ Z_{4(t-3)} + 0.005\ Z_{5(t-3)} - 0.023\ Z_{1(t-12)} - 0.052\ Z_{2(t-12)} + 0.029\ Z_{3(t-12)} + 0.02\ Z_{4(t-12)} - 0.024\ Z_{5(t-12)} + 0.042\ Z_{1(t-36)} + 0.048\ Z_{2(t-36)} + 0.752\ Z_{3(t-36)} + 0.035\ Z_{4(t-36)} + 0.066\ Z_{5(t-36)}$

While the result of the parameter estimation of the GSTAR ((1), 1,2,3,12,36) -SUR model with cross-covariance normalization location weights is as follows :

$\hat{Z}_{3t} = 0.033\ Z_{1(t-1)} + 0.003\ Z_{2(t-1)} + 0.407\ Z_{3(t-1)} + 0.023\ Z_{4(t-1)} + 0.014\ Z_{5(t-1)} + 0.018\ Z_{1(t-2)} + 0.078\ Z_{2(t-2)} - 0.348\ Z_{3(t-2)} + 0.019\ Z_{4(t-2)} + 0.069\ Z_{5(t-2)} + 0.076\ Z_{1(t-3)} - 0.061\ Z_{2(t-3)} - 0.028\ Z_{3(t-3)} + 0.026\ Z_{4(t-3)}) + 0\ Z_{5(t-3)} - 0.023\ Z_{1(t-12)} - 0.045\ Z_{2(t-12)} + 0.067\ Z_{3(t-12)}) + 0.017\ Z_{4(t-12)} - 0.022\ Z_{5(t-12)} + 0.035\ Z_{1(t-36)} + 0.042\ Z_{2(t-36)} + 0.744\ Z_{3(t-36)}) + 0.026\ Z_{4(t-36)} + 0.065\ Z_{5(t-36)}$

The result of parameter estimation of GSTAR ((1), 1,2,3,12,36) -SUR model with the weight of cross-correlation normalization location for Dau Subdistrict is as follows:

$\hat{Z}_{4t} = 0.033\ Z_{1(t-1)} + 0.009\ Z_{2(t-1)} - 0.009\ Z_{3(t-1)} + 0.253\ Z_{4(t-1)} + 0.016\ Z_{5(t-1)} + 0.023\ Z_{1(t-2)} + 0.089\ Z_{2(t-2)} + 0.086\ Z_{3(t-2)} + 0.13\ Z_{4(t-2)} + 0.069\ Z_{5(t-2)} + 0.076\ Z_{1(t-3)} - 0.075\ Z_{2(t-3)} + 0.038\ Z_{3(t-3)} + 0\ Z_{4(t-3)} + 0.005\ Z_{5(t-3)} - 0.023\ Z_{1(t-12)} - 0.057\ Z_{2(t-12)} - 0.01\ Z_{3(t-12)} - 0.119\ Z_{4(t-12)} - 0.023\ Z_{5(t-12)} + 0.042\ Z_{1(t-36)}) + 0.053\ Z_{2(t-36)} - 0.096\ Z_{3(t-36)} + 0.169\ Z_{4(t-36)} + 0.064\ Z_{5(t-36)}$

While the result of the parameter estimation of the GSTAR ((1), 1,2,3,12,36) -SUR model with cross-covariance normalization location weights is as follows :

$\hat{Z}_{4t} = 0.034\ Z_{1(t-1)} + 0.003\ Z_{2(t-1)} - 0.006\ Z_{3(t-1)} + 0.25\ Z_{4(t-1)} + 0.014\ Z_{5(t-1)} + 0.018\ Z_{1(t-2)} + 0.085\ Z_{2(t-2)} + 0.084\ Z_{3(t-2)} + 0.083\ Z_{4(t-2)} + 0.066\ Z_{5(t-2)} + 0.077\ Z_{1(t-3)} - 0.067\ Z_{2(t-3)} + 0.032\ Z_{3(t-3)} - 0.008\ Z_{4(t-3)} + 0\ Z_{5(t-3)} - 0.023\ Z_{1(t-12)} - 0.049\ Z_{2(t-12)} - 0.016\ Z_{3(t-12)} - 0.113\ Z_{4(t-12)} - 0.021\ Z_{5(t-12)} + 0.035\ Z_{1(t-36)} + 0.046\ Z_{2(t-36)}) - 0.087\ Z_{3(t-36)} + 0.203\ Z_{4(t-36)} + 0.063\ Z_{5(t-36)}$

The result of parameter estimation of GSTAR ((1), 1,2,3,12,36) -SUR model with the weight of cross-correlation normalization location for Wagir Subdistrict is as follows:

$\hat{Z}_{5t} = 0.031\ Z_{1(t-1)} + 0.009\ Z_{2(t-1)} - 0.01\ Z_{3(t-1)} + 0.023\ Z_{4(t-1)} + 0.243\ Z_{5(t-1)} + 0.022\ Z_{1(t-2)} + 0.092\ Z_{2(t-2)} + 0.092\ Z_{3(t-2)} + 0.01\ Z_{4(t-2)} + 0.07\ Z_{5(t-2)} + 0.072\ Z_{1(t-3)} - 0.078\ Z_{2(t-3)} + 0.04\ Z_{3(t-3)} + 0.024\ Z_{4(t-3)} + 0.048\ Z_{5(t-3)} - 0.022\ Z_{1(t-12)} - 0.059\ Z_{2(t-12)} - 0.011\ Z_{3(t-12)} + 0.02\ Z_{4(t-12)} + 0.047\ Z_{5(t-12)} + 0.04\ Z_{1(t-36)} + 0.055\ Z_{2(t-36)} - 0.102\ Z_{3(t-36)} + 0.035\ Z_{4(t-36)} + 0.206\ Z_{5(t-36)}$

While the result of the parameter estimation of the GSTAR ((1), 1,2,3,12,36) -SUR model with cross-covariance normalization location weights is as follows :

$\hat{Z}_{5t} = 0.032\ Z_{1(t-1)} + 0.003\ Z_{2(t-1)} - 0.007\ Z_{3(t-1)} + 0.023\ Z_{4(t-1)} + 0.252\ Z_{5(t-1)} + 0.017\ Z_{1(t-2)} + 0.088\ Z_{2(t-2)} + 0.089\ Z_{3(t-2)} + 0.018\ Z_{4(t-2)} + 0.071\ Z_{5(t-2)} + 0.072\ Z_{1(t-3)} - 0.07\ Z_{2(t-3)} + 0.034\ Z_{3(t-3)} + 0.026\ Z_{4(t-3)} + 0.061\ Z_{5(t-3)} - 0.022\ Z_{1(t-12)} - 0.051\ Z_{2(t-12)} - 0.017\ Z_{3(t-12)} + 0.017\ Z_{4(t-12)} + 0.042\ Z_{5(t-12)} + 0.033\ Z_{1(t-36)} + 0.047\ Z_{2(t-36)} - 0.093\ Z_{3(t-36)} + 0.026\ Z_{4(t-36)} + 0.204\ Z_{5(t-36)}$

The following is a plot to predict rainfall data in each location:
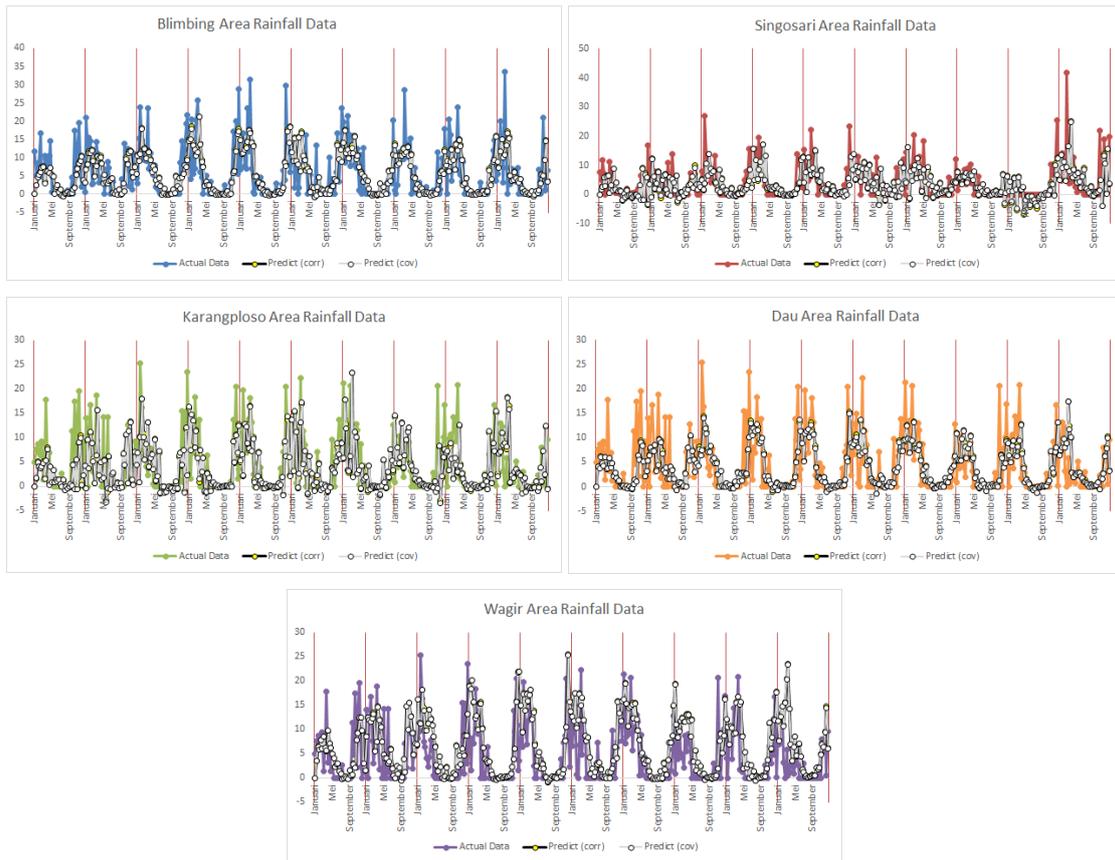
Figure 1. Actual Rainfall and Prediction of GSTAR ((1), 1,2,3,12,36) -SUR with the weight of Cross-Correlation and Cross-Covariance

Examination of the accuracy of the GSTAR ((1), 1,2,3,12,36) -SUR models was done by calculating the RMSE and R^2 prediction values in the model with the weighing location of the normalized cross-correlation and cross-covariance. The lower the RMSE value and the higher the R^2 prediction value, the better the accuracy of the GSTAR ((1), 1,2,3,12,36) -SUR model in generating the forecast value. Following is the examination of the accuracy of the GSTAR ((1), 1,2,3,12,36) -SUR models presented in Table 6:

Table 6. Accuracy Examination of GSTAR ((1), 1,2,3,12,36) -SUR

Location	Cross-Correlation Weight Model			Cross-Correlation Weight Model		
	RMSE Data Training	RMSE Data Testing	R^2 *prediction*	RMSE Data Training	RMSE Data Testing	R^2 *prediction*
Blimbing			0.579			0.558
Singosari			0.609			0.599
Karangploso	5.796	10.471	0.707	5.779	10.433	0.720
Dau			0.565			0.595
Wagir			0.328			0.336

Based on the results of the accuracy of the GSTAR ((1), 1,2,3,12,36) -SUR models in Table 6, the GSTAR ((1), 1,2,3,12,36) -SUR models that use correlation weights cross has RMSE training data value of 5.796 and RMSE on testing data is 10.471. Whereas in the GSTAR ((1), 1,2,3,12,36) -SUR model which uses the cross-covariance weight, model, the RMSE value of training data is 5,779 and RMSE testing data is 10,433. If the RMSE values

in the two models are compared, the RMSE values of the two models are relatively the same.

Besides being done by calculating the RMSE value, checking the accuracy of the model is also done by calculating the R2 prediction value at each location. As shown in Table 6, R^2 prediction values on GSTAR ((1), 1,2,3,12,36) -SUR models that use cross-covariance weights, are higher than R^2 prediction in models with cross-correlation weights, except in locations Blimbing and Singosari Districts.

CONCLUSIONS

The cross-covariance model produces a better level of accuracy in terms of lower RMSE values and higher R^2 values, especially for Karangploso, Dau, and Wagir areas.

ACKNOWLEDGMENTS

REFERENCES

[1]   P. E. Pfeifer and S. J. Deutsch, "Identification and interpretation of first-order space-time ARMA models," *Technometrics*, 1980.

[2]   P. E. Pfeifer and S. J. Deutsch, "Seasonal Space-Time ARIMA Modeling," *Geogr. Anal.*, 1981.

[3]   Borovkova, Lopuha, and B. N. Ruchjana, "Generalized S-TAR with Random Weights," in *Proceeding of the 17th International Workshop on Statistical Modeling*, 2002.

[4]   B. N. Ruchjana, "Study on the Weight Matrix in the Space-Time Autoregressive Model," in *Proceeding of the Tenth International Symposium on Applied Stochastic Models and Data Analysis (ASMDA)*, 2001, pp. 789–794.

[5]   B. N. Ruchjana, "Pemodelan Kurva Produksi Minyak Bumi Menggunakan Model Generalisasi STAR," Bogor, 2001.

[6]   A. Iriany, Suhariningsih, B. N. Ruchjana, and Setiawan, "Prediction of Precipitation Data at Batu Town Using the GSTAR (1,p) -SUR Model," *J. Basic Appl. Sci. Res.*, vol. 3, no. 6, pp. 860–865, 2013.

[7]   A. D. Sulistyono, W. H. Nugroho, R. Fitriani, and A. Iriany, "Hybrid Model GSTAR-SUR-NN For Precipitation Data," *Cauchy*, vol. 4, no. 2, p. 74, 2016.

[8]   A. D. Sulistyono, W. Hadi Nugroho, and A. Iriany, "Development of Hybrid Model GSTAR-SUR-NN and Application for Rainfall Forecasting," in *1st International Conference Pure Applied Resources Univ. Muhammadiyah Malang*, 2015, p. 104.

[9]   A. Iriany, W. M. Firdaus, W. H. Nugroho, and A. D. Sulistyono, "Rainfall Forecasting Using GSTAR-SUR-NN Approach in West Java Province," in *International Conference on Science, Engineering, Built Environment, and Social Science*, 2016, p. 1.

[10]   Suhartono and R. M. Atok, "Pemilihan Bobot Lokasi yang Optimal pada Model GSTAR," in *Prosiding Konferensi Nasional Matematika XIII*, 2006.

[11]   Suhartono and Subanar, "The Optimal Determination of Space Weight in GSTAR Model by using Cross-correlation Inference," *J. Quant. Methods*, vol. 2, no. 2, pp. 45–53, 2006.

[12]   A. D. Sulistyono, W. H. Nugroho, and A. Iriany, "Location Weight of GSTAR Model for Heterogeneity Variance of Precipitation Data," in *International Conference on*

Science, Engineering, Built Environment, and Social Science, 2016, p. 6.

[13] T. V. Apanasovich and M. G. Genton, "Cross-covariance functions for multivariate random fields based on latent dimensions," *Biometrika*, vol. 97, no. 1, pp. 15–30, 2010.

[14] S. Efromovich and E. Smirnova, "Statistical Analysis of Large Cross-Covariance and Cross-Correlation Matrices Produced by fMRI Images," *J. Biom. Biostat.*, vol. 05, no. 02, 2013.