Rainfall forecasting in agricultural areas using GSTAR-SUR model

To cite this article: A D Sulistyono et al 2020 IOP Conf. Ser.: Earth Environ. Sci. 458 012041

View the article online for updates and enhancements.
Rainfall forecasting in agricultural areas using GSTAR-SUR model

A D Sulistyono1*, Hartawati2, A Iriany2, N W Suryawardhani3 and A Iriany3

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

*E-mail: agusdwistat@ub.ac.id

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 normalization of cross-correlation. 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 2005 to December 2015. 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.

1. Introduction

The spatiotemporal model that had been developed firstly was the Space-Time Autoregressive (STAR) model developed by Pfeifer & Deutsch [1-2]. The Space-Time Autoregressive (STAR) model has 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 GSTAR model 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. To overcome data that have a seasonal pattern and not stationary, Iriany, et al [6] developed GSTAR-SUR model. Furthermore, the use of hybrid model GSTAR-SUR-NN was also developed for data that has a nonlinear pattern [7-9].

The location weights used in spatiotemporal models contribute to the level of model accuracy. The location weights that are often used are uniform weight, inverse distance, and normalization cross-correlation [10-11]. The weight of the location considers the proximity between locations. The use of
the location weights mentioned above is less relevant for data that has a high level of variability. Therefore, we need the 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 normalization cross-covariance weight has been applied in the research of Apanasovich, et al [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.

2. Materials and Methods
The data period used is January 2005 to December 2018, where data for conducting a training (in-sample) is data from January 2005 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 the equation

\[ w_r = \frac{r_r(1)}{\sum_{k \neq i} r_k(1)} \]  

[11]

and the normalization weight of cross-covariance

\[ \frac{\sum_{i=1}^{n} [z_i(t) - \bar{z}] [z_i(t-k) - \bar{z}]^T}{\sqrt{\sum_{i=1}^{n} [z_i(t) - \bar{z}]^2} \sqrt{\sum_{i=1}^{n} [z_i(t-k) - \bar{z}]^2}} \]  

[13], [14]. The next process is to do a GSTAR-OLS analysis to get the residual value with the equation

\[ Z(t) = \mathbf{\mu}(t) + [\mathbf{\Phi}_{01} + \mathbf{\Phi}_{11} W] Z(t-1) + \epsilon(t) \].

Next, calculate the \( \text{var}(\epsilon) = \mathbf{\Omega} \) matrix with the equation

\[ \mathbf{\Sigma} = \begin{bmatrix}
\sigma_{11} & \sigma_{12} & \ldots & \sigma_{1m} \\
\sigma_{21} & \sigma_{22} & \ldots & \sigma_{2m} \\
\vdots & \vdots & \ddots & \vdots \\
\sigma_{m1} & \sigma_{m2} & \ldots & \sigma_{mm}
\end{bmatrix} \]

The next step is estimating the GSTAR (1, p)-SUR parameter using the formula

\[ \mathbf{\beta} = (\mathbf{X}'\mathbf{\Omega}^{-1}\mathbf{X})^{-1}\mathbf{X}'\mathbf{\Omega}^{-1}\mathbf{y} \].

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

3. Results and Discussion
This research was carried out by taking daily rainfall data obtained from the rain heading point for the Blimming, 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>
<thead>
<tr>
<th>Location</th>
<th>N</th>
<th>Average (mm)</th>
<th>Standard Deviation (mm)</th>
<th>Minimum (mm)</th>
<th>Maximum (mm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Blimming</td>
<td>360</td>
<td>5.682</td>
<td>6.909</td>
<td>0</td>
<td>33.5</td>
</tr>
<tr>
<td>Singosari</td>
<td>360</td>
<td>3.93</td>
<td>5.575</td>
<td>0</td>
<td>41.75</td>
</tr>
<tr>
<td>Karangploso</td>
<td>360</td>
<td>4.302</td>
<td>5.71</td>
<td>0</td>
<td>25.36</td>
</tr>
<tr>
<td>Dau</td>
<td>360</td>
<td>4.564</td>
<td>8.187</td>
<td>0</td>
<td>43.63</td>
</tr>
<tr>
<td>Wagir</td>
<td>360</td>
<td>7.08</td>
<td>8.187</td>
<td>0</td>
<td></td>
</tr>
</tbody>
</table>
Based on Table 1 above, it is shown that the average rainfall data in Wagir District was the highest and Singosari District has the lowest average rainfall. The standard deviation on each location was greater than the average, indicating a high level of rainfall variation. In addition, the heterogeneity of the observation location can be measured by calculating the Gini Index. The higher the index value, the higher heterogeneity of rainfall data in all the locations will be. The Gini Index can be counted as follow:

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

\[ = 1 + \frac{1}{360} - \frac{2}{360^2} \cdot 9200.68 \]

\[ = 0.975 \]

Based on the results of the Gini index calculation, the high 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 variety of rainfall in each location.

<table>
<thead>
<tr>
<th>Location</th>
<th>( \lambda )</th>
<th>Transformation 1</th>
<th>Transformation</th>
<th>( \lambda )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Blimbing</td>
<td>- 0.27</td>
<td>Z(^{0.27})</td>
<td>1.0</td>
<td></td>
</tr>
<tr>
<td>Singosari</td>
<td>- 0.68</td>
<td>Z(^{0.68})</td>
<td>1.0</td>
<td></td>
</tr>
<tr>
<td>Karangploso</td>
<td>- 0.50</td>
<td>Z(^{0.5})</td>
<td>1.0</td>
<td></td>
</tr>
<tr>
<td>Dau</td>
<td>- 0.50</td>
<td>Z(^{0.5})</td>
<td>1.0</td>
<td></td>
</tr>
<tr>
<td>Wagir</td>
<td>- 0.19</td>
<td>Z(^{0.19})</td>
<td>1.0</td>
<td></td>
</tr>
</tbody>
</table>

Based on Table 4.3 the initial \( \lambda \) 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. The results of the ADF test were presented in Table 2.

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 3. Stationary testing on the average rainfall of each location.

<table>
<thead>
<tr>
<th>Location</th>
<th>t-Statistics</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Blimbing</td>
<td>-5.974</td>
<td>0.000</td>
</tr>
<tr>
<td>Singosari</td>
<td>-5.556</td>
<td>0.000</td>
</tr>
<tr>
<td>Karangploso</td>
<td>-7.831</td>
<td>0.000</td>
</tr>
<tr>
<td>Dau</td>
<td>-7.215</td>
<td>0.000</td>
</tr>
<tr>
<td>Wagir</td>
<td>-7.425</td>
<td>0.000</td>
</tr>
</tbody>
</table>

Table 4. Matrix Partial Autocorrelation Scheme (MPACF).

<table>
<thead>
<tr>
<th>Variable/Lag</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
</tr>
</thead>
<tbody>
<tr>
<td>BLIMING</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SINGOSARI</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>KARANGPLOSO</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DAU</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>WAGIR</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

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.

<table>
<thead>
<tr>
<th>Order</th>
<th>AIC</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>14.67906</td>
</tr>
<tr>
<td>2</td>
<td>14.52868</td>
</tr>
<tr>
<td>3</td>
<td>14.52445</td>
</tr>
</tbody>
</table>

Based on the results of the accuracy-test using 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.

The identification of the GSTAR model is also done by univariate ACF and PACF plots at each location. Based on the ACF plot, the rainfall data for each location is indicated by seasonal patterns. 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 location $i$ are 4 locations so that the cross-correlation normalization matrix is as follows:
While the magnitude of cross-covariance normalization weighting based on calculations is as follows:

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

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:

\[
\hat{Z}_{1t} = 0.116 Z_{1(t-1)} + 0.003 Z_{2(t-1)} - 0.006 Z_{3(t-1)} + 0.022 Z_{4(t-1)} + 0.015 Z_{5(t-1)} + 0.212 Z_{1(t-2)} + 0.083 Z_{2(t-2)} + 0.081 Z_{3(t-2)} + 0.018 Z_{4(t-2)} + 0.072 Z_{5(t-2)} - 0.092 Z_{1(t-3)} + 0.066 Z_{2(t-3)} + 0.031 Z_{3(t-3)} + 0.025 Z_{4(t-3)} + 0.004 Z_{5(t-3)} + 0.06 Z_{1(t-12)} - 0.048 Z_{2(t-12)} - 0.016 Z_{3(t-12)} + 0.016 Z_{4(t-12)} - 0.023 Z_{5(t-12)} + 0.256 Z_{1(t-36)} + 0.045 Z_{2(t-36)} - 0.085 Z_{3(t-36)} + 0.025 Z_{4(t-36)} + 0.068 Z_{5(t-36)}
\]

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:

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

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:

\[
\hat{Z}_{1t} = 0.024 Z_{1(t-1)} + 0.206 Z_{2(t-1)} - 0.006 Z_{3(t-1)} + 0.021 Z_{4(t-1)} + 0.01 Z_{5(t-1)} + 0.013 Z_{1(t-2)} - 0.281 Z_{2(t-2)} + 0.081 Z_{3(t-2)} + 0.016 Z_{4(t-2)} + 0.05 Z_{5(t-2)} + 0.055 Z_{1(t-3)} + 0.639 Z_{2(t-3)} + 0.03 Z_{3(t-3)} + 0.023 Z_{4(t-3)} + 0.017 Z_{1(t-12)} + 0.301 Z_{2(t-12)} - 0.015 Z_{3(t-12)} + 0.015 Z_{4(t-12)} - 0.016 Z_{5(t-12)} + 0.025 Z_{1(t-36)} - 0.04 Z_{2(t-36)} - 0.084 Z_{3(t-36)} + 0.023 Z_{4(t-36)} + 0.047 Z_{5(t-36)}
\]

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:

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

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:

\[
\hat{Z}_{1t} = 0.033 Z_{1(t-1)} + 0.003 Z_{2(t-1)} + 0.407 Z_{3(t-1)} + 0.023 Z_{4(t-1)} + 0.014 Z_{5(t-1)} + 0.018 Z_{1(t-2)} + 0.078 Z_{2(t-2)} - 0.348 Z_{3(t-2)} + 0.019 Z_{4(t-2)} + 0.069 Z_{5(t-2)} + 0.076 Z_{1(t-3)} - 0.061 Z_{2(t-3)} - 0.028 Z_{3(t-3)} + 0.026 Z_{4(t-3)} + 0.0 Z_{5(t-3)} - 0.023 Z_{1(t-12)} - 0.045 Z_{2(t-12)} + 0.067 Z_{3(t-12)} + 0.017 Z_{4(t-12)} - 0.022 Z_{5(t-12)} + 0.035 Z_{1(t-36)} + 0.042 Z_{2(t-36)} + 0.744 Z_{3(t-36)} + 0.026 Z_{4(t-36)} + 0.065 Z_{5(t-36)}
\]
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:

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

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

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

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:

\[ \hat{Z}_w = 0.031 Z_{4(t-1)} + 0.009 Z_{2(t-1)} - 0.01 Z_{3(t-1)} + 0.023 Z_{4(t-1)} + 0.243 Z_{5(t-1)} + 0.022 Z_{1(t-2)} + 0.092 Z_{2(t-2)} + 0.104 Z_{3(t-2)} + 0.072 Z_{4(t-2)} + 0.072 Z_{5(t-2)} + 0.072 Z_{1(t-3)} + 0.078 Z_{2(t-3)} + 0.04 Z_{3(t-3)} + 0.024 Z_{4(t-3)} + 0.048 Z_{5(t-3)} + 0.022 Z_{3(t-12)} - 0.059 Z_{2(t-12)} - 0.011 Z_{3(t-12)} + 0.02 Z_{4(t-12)} + 0.044 Z_{5(t-12)} + 0.04 Z_{1(t-36)} + 0.055 Z_{2(t-36)} + 0.102 Z_{3(t-36)} + 0.035 Z_{4(t-36)} + 0.206 Z_{5(t-36)} \]

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

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

Examination of the accuracy of the GSTAR ((1), 1,2,3,12,36) -SUR models was done by calculating the RMSE and R² 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 R² 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>
<thead>
<tr>
<th>Location</th>
<th>RMSE Data Training</th>
<th>RMSE Data Testing</th>
<th>R² prediction</th>
<th>RMSE Data Training</th>
<th>RMSE Data Testing</th>
<th>R² prediction</th>
</tr>
</thead>
<tbody>
<tr>
<td>Blimbing</td>
<td>5.796</td>
<td>10.471</td>
<td>0.579</td>
<td>5.779</td>
<td>10.433</td>
<td>0.558</td>
</tr>
<tr>
<td>Singosari</td>
<td>5.152</td>
<td>8.609</td>
<td>0.609</td>
<td>5.152</td>
<td>8.609</td>
<td>0.609</td>
</tr>
<tr>
<td>Karangploso</td>
<td>5.328</td>
<td>10.354</td>
<td>0.307</td>
<td>5.328</td>
<td>10.354</td>
<td>0.307</td>
</tr>
<tr>
<td>Dau</td>
<td>5.065</td>
<td>7.092</td>
<td>0.552</td>
<td>5.065</td>
<td>7.092</td>
<td>0.552</td>
</tr>
<tr>
<td>Wagir</td>
<td>5.796</td>
<td>10.471</td>
<td>0.579</td>
<td>5.779</td>
<td>10.433</td>
<td>0.558</td>
</tr>
</tbody>
</table>

The following is a plot to predict rainfall data in each location:
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, R² prediction values on GSTAR ((1), 1,2,3,12,36) -SUR models that use cross-covariance weights, are higher than R² prediction in models with cross-correlation weights, except in locations Blimbing and Singosari Districts.

4. Conclusion
The cross-covariance model produces a better level of accuracy in terms of lower RMSE values and higher R² values, especially for Karangploso, Dau, and Wagir areas.

Acknowledgments
We would like to thank to the University of Muhammadiyah Malang and Brawijaya University for funding and support of this research.

References
France) 789–794.


