Modeling the Flood Disaster in South Kalimantan Using Geographically Weighted Regression and Mixed Geographically Weighted Regression

Yuniar Farida[1*], *Monika* Refiana Nurfadila[1], *Putroue* Keumala Intan[1], *Hani* Khaulasari[1], *Nurissaidah* Ulinnuha[1], *Wika* Dianita Utami[1], and *Dian* Yuliati[1]

[1] Department of Mathematics, Faculty of Science and Technology, UIN Sunan Ampel Surabaya, East Java, Indonesia

Abstract. The flood disaster in South Kalimantan is a crucial problem that needs to be addressed because the impact is relatively severe. So, this study aims to model flood disasters in South Kalimantan based on factors suspected to be the cause, including population density, rainfall, residential area, and forest area. This study uses two methods of spatial statistics, namely the Geographically Weighted Regression (GWR) and Mixed Geographically Weighted Regression (MGWR) methods. The weighting used is Adaptive Gaussian. The modeling results show that the GWR model is superior in explaining the causes of flood events in South Kalimantan, which is indicated by the highest coefficient of determination value of 95.62% compared to the regression and MGWR models. Nonetheless, the MGWR model can explain the causes of flooding in Kalimantan. The GWR and MGWR models show that the area that is vulnerable to flooding is Balangan District. The results of this study contribute to providing alternative information for disaster mitigation to minimize losses.

1 Introduction

Regarding geography, climatology, and demographics, Indonesia is a disaster-prone area. From a geographical perspective, Indonesia is located at the junction of three main plates: Indo-Australian, East Pacific, and Eurasian [1]. Climatologically, Indonesia is also in a tropical climate zone with relatively high rainfall [2]. It was supported and strengthened by data from the World Risk Report 2018, which stated that Indonesia was ranked 36th out of 172 countries most prone to natural disasters [3]. Regarding demographics, Indonesia is the fourth most populous country in the world, and it is predicted that population growth in the next 25 years will still be high [4].

Natural disasters result in severe suffering for humans, such as loss of property, human casualties, reduced natural resources, and dynamic environmental changes [5]. One of the sectors of life that was greatly affected by natural disasters was the economy because business actors suffered losses due to financial decline [6]. Natural disasters that often occur in Indonesia are geological natural disasters and hydrometeorological natural disasters.

* Corresponding author: yuniarfarida@uinsby.ac.id

Geological natural disasters have the characteristics of occurring on earth, for example, earthquakes, volcanic eruptions, and tsunamis. Meanwhile, hydrometeorological natural disasters are closely related to climate change, which causes natural disasters such as floods, tornadoes, and landslides. Hydrometeorological disaster events in Indonesia occurred at 97.12%, one dominated by floods [7].

One of the areas of Indonesia that is often hit by floods is the Province of South Kalimantan. It is noted that flood cases in South Kalimantan were 334 times higher than in other Kalimantan provinces [8]. Based on the publication of the Regional Disaster Management Agency (BPBD), from 2018 to 2020, flood disasters in South Kalimantan increased; in 2018, there were 30 cases of flooding. Meanwhile, in 2019, there were 66 cases of floods, and in 2020, it became the worst period of floods, namely 349 instances of floods. In addition, in mid-January 2021, a flood occurred which hit ten regencies/cities in South Kalimantan, including Banjarbaru City, Banjar Regency, Hulu Sungai Selatan Regency, Hulu Sungai Tengah Regency, Tapin Regency, Tanah Laut City, Balangan Regency, Tabalong Regency and Banjarmasin City [9]. Therefore, this study aims to model flood cases in South Kalimantan based on the causal factors.

From several previous studies, several factors have contributed to the occurrence of floods in various areas, including high rainfall causing rivers to overflow [10], reduced forest land is also a trigger for flooding, forest fires causing lack of forest land [11] and the expansion of new land used for coal mining and plantation land [12]. The triggering factor for flooding is also caused by an increase in population, which is increasing the demand for large residential land, where the construction of large residential land will disrupt water catchment areas [13]. Thus, flooding is a complex event due to the interaction of environmental, social variables, and natural conditions [14]. Socio-environmental variables include all community activities that pay little attention to nature. In contrast, biological variables are related to geographical constraints, referred to as spatial effects in the statistical concept. In this case, the spatial impact reflects the pattern of interrelationships between regions with a specific functional relationship between the explanatory and response variables. Therefore, the method used to identify flood-prone areas in South Kalimantan can use spatial statistical methods.

One spatial statistical method often used in research is Geographically Weighted Regression (GWR). GWR is developing a multiple linear regression model by assigning spatial positions to the regression parameters [15]. The regression coefficients include the spatial distance to the target point and the covariates. GWR allows each data point to be weighted based on its distance from the regression point, so data points closer to the regression point are heavily weighted in the local regression than data points that are more distant [16]. Previous research using the GWR method, such as research by Purwaningsih et al. [17] modeling flood cases in Central Java, showed that the highest flood events are in the Cilacap area, where in the Cilacap area, flood cases are caused by rainy days and the room with a coefficient of determination of 0.69. Chen et al. [18] analyzed the prevalence of tuberculosis cases in Kashgar, China, by obtaining a coefficient of determination of 0.7452. Wu et al. [19] compared the GWR and other spatial statistical methods. The results show that GWR predicts with higher accuracy than other spatial methods.

Furthermore, another Spatial Statistics method, the Mixed Geographically Weighted Regression (MGWR) method, modifies the GWR. The MGRW was first developed by determining that some regression coefficients are globally constant while others vary geographically [20]. This modification increases the model's flexibility for a broader range of situations by considering the spatial autocorrelation of model residuals [21]. The emergence of the MGWR method was due to the analysis of the regression method only estimating the average parameters in the entire region. At the same time, the MGWR method also pays attention to local parameters. So, objectively, the MGWR method compares the combination model between fixed coefficients and varying or local coefficients [22]. Several previous studies using the MGWR method, such as research by Shabrina et al. [23] comparing the GWR and MGWR models on short-term rental platforms in the urban tourism sector, obtaining variables that significantly influence globally and locally with a coefficient of determination of 0.767. Bera and Kangalli [24] analyzed the global and local factors on office rent in Istanbul with a coefficient of determination of 0.77. Chao et al. [25] compared GWR and MGWR in predicting rainfall and found that MGWR had the best performance with a coefficient of determination of 0.863.

Reviewing several previous studies shows that the GWR and MGWR models in various case studies have outperformed each other. So, in this study, to model flood-prone areas and identify the factors that influence flood cases in South Kalimantan, use these two methods, namely GRW and MGWR, and compare the performance of the two models in modeling flood cases in South Kalimantan. This research provides a valuable mapping of disaster-prone areas so that disaster prevention and management mitigation measures can be directed more effectively.

2 Methods

This study analyzes the factors that influence the occurrence of flooding in South Kalimantan through several stages, which will be shown in the research flow in Figure 1:

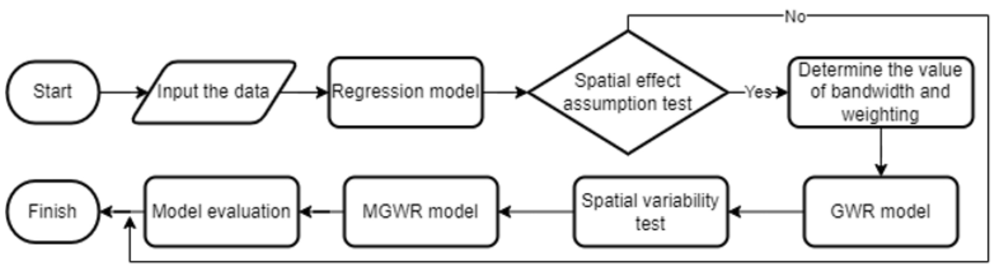

Fig 1. Research Flowchart

2.1 Data

This study used secondary data, namely flood events in every district/city in South Kalimantan Province in 2021, which can be accessed on the BPBD South Kalimantan website. Population density and forest area data can be accessed at BPS Kalimantan Selatan. Meanwhile, rainfall variable data is accessed on the BMKG online page, and data on the location of residential areas comes from the South Kalimantan Residential Service. The variable data with its unit and the sample data are shown in Table 1 and Table 2:

Table 1. Research Data

Symbol	Variable	Unit
y	Number of flood disaster	Incident
x_1	Population density	Soul/km^2
x_2	Rainfall	Millimeter
x_3	Residential area	Hectare

Symbol	Variable	Unit
x_4	Forest area	Hectare
u	Latitude	Degrees
v	Longitude	Degrees

Table 2. Sample Research Data

Regency/City	y	x_1	x_2	x_3	x_4	u	v
Regency.Balangan	89	70.39	762.133	27.31	90966.39	-2.5275	115.7302
Regency. Banjar	53	122.56	836.05	322.98	142058.1	-3.7287	115.5959
Banjarbaru City	21	697.45	836.05	138.7	2018.69	-3.5694	114.9224
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
Regency.Tanah Laut	37	97.58	836.05	1003.51	105829	-4.1785	115.3736
Regency. Tapin	4	71.02	614.133	37.37	27110.31	-3.1986	115.5074

2.2 Linear regression

Linear regression is a statistical method used to determine the relationship between the dependent and independent variables [26], forming a linear curve. In addition, the linear regression model can be used to predict the independent variables [27]. The linear regression model can be written as follows [28]:

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \cdots + \beta_p x_{ip} + \varepsilon_i \tag{1}$$

Where y is the dependent variable and the value $i = 1,2, \ldots, n$ with $x_1, x_2, \ldots, x_p$ is the independent variable of some p variables. Meanwhile, to estimate the parameters of the linear regression model using Ordinary Least Square (OLS) by minimizing the remaining squared residuals in the regression model as follows [29]:

$$\hat{\beta} = (X^T X)^{-1} X^T y \tag{2}$$

2.3 Geographically weighted regression

In modeling using the GWR method, it is necessary to analyze spatial heterogeneity or differences in conditions between regions regarding geography, socio-cultural, and others. By using the Breusch-Pagan test statistics below:

$$BP = \frac{1}{2}[f^T Z (Z^T Z)^{-1} Z^T f] \sim \chi_p^2 \tag{3}$$

Spatial heterogeneity is detected when the value $BP > \chi_{\alpha,p}^2$ or when the p-value is less than the significance value ($\alpha = 0{,}05$). Next is determining the weighting function representing the observation area data [28]. The weighting function used in the GWR model is the Adaptive Gaussian kernel function, written as follows.

$$w_j(u_i, v_i) = exp\left(-\frac{1}{2}\left(\frac{d_{ij}}{h_{i(p)}}\right)^2\right) \tag{4}$$

Where $h_{i(p)}$ is the i-th bandwidth value, and d_{ij} is the distance between the i-th location and the j-th location. In determining the distance between sites using the Euclidean distance equation as follows.

$$d_{ij} = \sqrt{(u_i - u_j)^2 + (v_i - v_j)^2} \tag{5}$$

Bandwidth determination is used to find out the limitations related to the extent of the connection between the points of the observation area and other areas. Bandwidth selection is by choosing the maximum based on the approach using cross-validation (CV). The validation statistical test is as follows [30]:

$$CV = \sum_{i=1}^{n} \left(y_i - \hat{y}_{\neq 1}(h)\right)^2 \tag{6}$$

Where $\hat{y}_{\neq 1}(h)$ represents the projected value y_i on the i-th observation, and h is the bandwidth value. When the maximum bandwidth value is obtained, the cross-validation value is minimal. Besides that, statistically, the GWR equation can be written as follows:

$$y_i = \beta_0(u_i, v_i) + \sum_{k=1}^{p} \beta_k(u_i, v_i)x_{ik} + \varepsilon_i\,, \quad i = 1,2 \ldots., n \tag{7}$$

Where y_i is the observed value of the i-bound variable, x_{ik} is the observed value of the k-independent variable in the i-th observation. Meanwhile, β is the parameter vector and (u_i, v_i) is the coordinate point of region i and ε_i is the i-th error.

2.4 Mixed geographically weighted regression

MGWR combines the linear regression method with weighted regression, where the independent variables used do not all have a global effect on all observation points. Still, some have an impact locally [31]. The MGWR model can be written in the following equation:

$$\mathrm{y_i} = \beta_0(\mathrm{u_i}, \mathrm{v_i}) + \sum_{\mathrm{k}=1}^{\mathrm{q}} \beta_{\mathrm{k}}(\mathrm{u_i}, \mathrm{v_i})\mathrm{x_{ik}} + \sum_{\mathrm{k}=\mathrm{q}+1}^{\mathrm{p}} \beta_{\mathrm{k}}\mathrm{x_{ik}} + \varepsilon_{\mathrm{i}} \tag{8}$$

Estimating the parameters in the MGWR model through two steps to estimate the global coefficient using the Ordinary Least Square method is shown below.

$$\hat{\beta}_p = \left(\beta_1, \beta_2, \cdots, \beta_{p_q}\right)^T = \left[X_p^T(I - S_q)^T(I - S_q)X_p\right]^{-1} X_p^T(I - S_q)^T(I - S_q)Y \tag{9}$$

where:

$$S_q = \begin{pmatrix} X_{q1}^T\left[W(u_1, v_1)X_q\right]^{-1} X_q^T W(u_1, v_1) \\ X_{q2}^T\left[W(u_2, v_2)X_q\right]^{-1} X_q^T W(u_2, v_2) \\ \vdots \\ X_{qn}^T\left[W(u_n, v_n)X_q\right]^{-1} X_q^T W(u_n, v_n) \end{pmatrix}$$

$W(u_i, v_i)$ was obtained from $W(u_i, v_i) = diag\,[w_1(u_i, v_i), w_2(u_i, v_i), \cdots, w_n(u_i, v_i)]$. Meanwhile, to estimate the local coefficients, use Weighted Least Square as follows.

$$\begin{aligned} \hat{\beta}_q(u_i, v_i) &= \left(\beta_1(u_i, v_i), \beta_2(u_i, v_i), \cdots, \beta_{p_q}(u_i, v_i)\right)^T \\ &= \left[X_q^T W\,(u_i, v_i)X_q\right]^{-1} X_q^T W(u_i, v_i)\left(Y - X_p\hat{\beta}_p\right) \end{aligned} \tag{10}$$

2.5 Model evaluation

To determine the best model that explains the relationship between the independent and dependent variables by comparing each coefficient of determination (R^2). The following statistical tests are used to calculate the coefficient of determination [32]:

$$R^2 = \frac{\sum_i^n (\hat{y}_i - \bar{y})^2}{\sum_i^n (y_i - \bar{y})^2} \times 100 \quad (11)$$

For y_i is the data from the actual dependent variable, $\hat{y}_i$ which is the result of the estimation of subject i, and for$\bar{y}$ is the average of the actual dependent variable data.

3 Result and discussion

The data used in this study shows that the average number of flood events in South Kalimantan in 2021 is 26. The highest flood events, 89 floods, occurred in the Balangan Regency area, and the lowest flood events occurred in the Kotabaru Regency area, which did not experience flooding in 2021.

3.1 Linear regression models

The modeling results between the dependent and independent variables using linear regression show that the whole or part test is insignificant at α = 0.05. However, this model can still be used because the four independent variables have an effect, although not substantial. Meanwhile, the classical assumption test results include normal distribution, no multicollinearity detected, heteroscedasticity detected, and no autocorrelation detected. So, the model of linear regression can be written as follows:

$$Y = -8.110 - 9.379\, x_1 + 0.1727\, x_2 - 51.73 x_3 - 5.173 \times 10^{-5} x_4$$

3.2 Geographically weighted regression models

From the spatial heterogeneity test results, it was obtained that the Breusch-Pagan test value was as significant as 10.71 > $\chi^2_{0.05;4}$ = 9.487729 so that spatial heterogeneity was detected. In addition, the maximum bandwidth was obtained for each observation area with a cross-validation value of 139518.4. A weighting value will be obtained from acquiring the maximum bandwidth, which will be used in estimating the GWR parameters. Table 3 shows the results of the conformity test of the GWR parameters based on the hypothesis:
$H_0 : \beta_k(u_i, v_i) = \beta_k,\ k = 0,1,2,\dots,p \ and\ i = 1,2,\dots,n$
H_1 : There is at least one $\beta_k(u_i, v_i) \neq \beta_k$

Table 3. Conformity of the GWR Model

Sources of Residual Diversity	The sum of Residual Squares	F_G
Regression	4684.7999	15.139
GWR	309.4494	

Based on Table 3, the value obtained from the conformity test of the GWR model for the Adaptive Gaussian weighting is $F_G > F_{0,05;2.051,8} = 4{,}4321$. So, it can be concluded that the GWR model significantly differs from the regression model. The flood case model using the GWR method is presented in Table 4:

Table 4. GWR Models

Regency/City	Model
Regency. Balangan	$\hat{y}_1 = -594.7438 + 0.96850\, x_2 - 0.00072\, x_4$
Regency. Banjar	$\hat{y}_2 = -77.36873$
Banjarbaru City	$\hat{y}_3 = -33.94032$
Banjarmasin City	$\hat{y}_4 = -31.38779$
Regency. Barito Kuala	$\hat{y}_5 = -30.68241$
Regency. Hulu Sungai Selatan	$\hat{y}_6 = -343.92335 + 0.59519\, x_2 - 0.00039\, x_4$
Regency. Hulu Sungai Tengah	$\hat{y}_7 = -501.18453 + 0.83780\, x_2 - 0.00062\, x_4$
Regency. Hulu Sungai Utara	$\hat{y}_8 = -274.10461 + 0.50258\, x_2 - 0.00033\, x_4$
Regency. Kotabaru	$\hat{y}_9 = -91.47201 + 0.17592\, x_2 - 0.00006\, x_4$
Regency. Tabalong	$\hat{y}_{10} = -415.67264 + 0.72580\, x_2 - 0.00056\, x_4$
Regency. Tanah Bumbu	$\hat{y}_{11} = -88.50047 + 0.165564 x_2$
Regency. Tanah Laut	$\hat{y}_{12} = -64.37234 + 0.11699\, x_2$
Regency. Tapin	$\hat{y}_{13} = -84.50295 + 0.15876 x_2$

The GWR model that is formed differs in each observation area based on factors significant to flood events, as shown in Table 4. The GWR model obtained has an effect only locally, while looking at variables that have a significant impact globally and locally can be seen from the results of the spatial variability in Table 5.

Table 5. Spatial Variability

	β_0	β_1	β_2	β_3	β_4
df_1	5.84494	3.40418	5.69521	8.20668	3.72527
df_2	4.24740	4.24740	4.24740	4.2474	4.24740
F_s	8.97017	0.90104	11.14407	1.04996	10.09491
$F_{\alpha;df_1;df_2}$	5.79324	6.12287	5.80629	5.64527	6.05891
p-values	0.02262	0.52232	0.01509	0.47285	0.01702

Based on Table 5, β_0 is a fixed constant value while for $\beta_1, \beta_2, \beta_3$ and β_4 is the coefficient value of each variable x_1, x_2, x_3 and x_4. The spatial variability test shows that local variables are based on a value of $F_s > F_{\alpha;df_1;df_2}$ or p-value < 0.05. In the rainfall variable (x_2), the value $F_s > F_{0.05;5.69521;4.24740} = 5.80629$ of the rainfall variable is identified as local variable. Likewise, the forest area variable (x_4) is included in the local variable because the value $F_s > F_{0.05;3.72527;4.24740} = 6.05891$. Meanwhile, variables that do not meet the requirements are included in the overall or global variables, namely population density (x_1) and residential area (x_3). Due to global and local variables, flood events in South Kalimantan can be modeled using the MGWR.

3.3 Mixed geographically weighted regression models

After the parameter results are obtained, the MGWR model suitability test is carried out based on the following hypotheses, and the test results are presented in Table 6.
$H_0 : \beta_k(u_i, v_i) = \beta_k,\ k = 0,1,2, \dots, p \ and\ i = 1,2, \dots, n$
H_1 : There is at least one $\beta_k(u_i, v_i) \neq \beta_k$

Table 6. Conformity of the MGWR Model

Sources of Residual Diversity	The sum of Residual Squares	F_M
Regression	4684.7999	1.3289707
MGWR	808.7917	

Based on Table 6, the results obtained from the conformity of the MGWR model that the value of $F_M < F_{0.05;10.47;4.43} = 4.973325$. Then, a decision can be taken to accept H_0 The MGWR model does not differ significantly from the regression model. Then, partial global and local parameter testing was carried out with the results presented in Table 7.

Table 7. Global Parameter Test

Parameter	Estimation	T_{phit}	p-values
β_1	0.002916	1.218649	0.244631
β_3	0.055787	3.382372	0.004906

Based on Table 7, the selected variables that significantly affect flood events in South Kalimantan are based on the value of $|T_{phit}| > t_{0.025;13.26} = 1.768264$. So, the variable that significantly influences South Kalimantan's flooding is the residential area. Furthermore, parameter testing in each region or locally is shown in Table 8.

Table 8. Local Parameter Test

Regency/City	$\boldsymbol{\beta_2}$	$\boldsymbol{t_{qhit(\beta_2)}}$	$\boldsymbol{\beta_4}$	$\boldsymbol{t_{qhit(\beta_4)}}$
Regency. Balangan	0.1035027848	8.7372511	-0.0003047981	-5.1503931
Regency. Banjar	2.126705×10^{-3}	0.09913697	5.153477×10^{-5}	0.9256725
Banjarbaru City	8.799135×10^{-3}	0.083183904	-2.446445×10^{-4}	-0.8456088
Banjarmasin City	-5.846312×10^{-4}	-0.01814085	2.035596×10^{-3}	0.2322195
Regency. Barito Kuala	4.594756×10^{-3}	0.14864034	9.557459×10^{-4}	0.1126934
Regency. Hulu Sungai Selatan	-9.631833×10^{-3}	-0.26229069	2.002192×10^{-4}	0.3222046
Regency. Hulu Sungai Tengah	-0.1886565390	-4.867264459	0.0018372292	6.6095992
Regency. Hulu Sungai Utara	8.553870×10^{-3}	0.51424459	-3.394986×10^{-5}	-0.4973051
Regency. Kotabaru	-2.717567×10^{-3}	-1.04789482	-1.759931×10^{-5}	-1.0478948
Regency. Tabalong	0.1841675780	10.55734672	-0.0005926734	-7.8606061
Regency. Tanah Bumbu	5.907381×10^{-2}	1.92546560	-1.088704×10^{-4}	-2.00039120
Regency. Tanah Laut	-3.680346×10^{-2}	-1.80243473	1.346684×10^{-4}	-2.9978848
Regency. Tapin	2.251715×10^{-2}	0.36521023	-4.478193×10^{-4}	-0.3777979

Based on Table 8, to conclude which areas are locally affected by rainfall and forest area, that is based on the value of $|T_{qhit}| > t_{0.025;13.26} = 1.768264$. After knowing which variables have global and local influences, the MGWR model can be formed, as shown in Table 9.

Table 9. MGWR Models

Regency/City	Model
Regency. Balangan	$\hat{y}_1 = 0.00157 + 0.10350\, x_2 + 0.55787 x_3 - 0.00030\, x_4$
Regency. Banjar	$\hat{y}_2 = 2.39998 \times 10^{-6} + 0.05578 x_3$
Banjarbaru City	$\hat{y}_3 = 1.21832 \times 10^{-5} + 0.05578 x_3$
Banjarmasin City	$\hat{y}_4 = -9.44818 \times 10^{-7} + 0.05578\, x_3$
Regency. Barito Kuala	$\hat{y}_5 = 6.32669 \times 10^{-6} + 0.05578\, x_3$
Regency. Hulu Sungai Selatan	$\hat{y}_6 = -1.56836 \times 10^{-5} + 0.05578 x_3$

Regency/City	Model
Regency. Hulu Sungai Tengah	$\hat{y}_7 = -0.00023 - 0.18865\, x_2 + 0.05578x_3 + 0.00183x_4$
Regency. Hulu Sungai Utara	$\hat{y}_8 = 1.46802 \times 10^{-5} + 0.05578x_3$
Regency. Kotabaru	$\hat{y}_9 = -3.56574 \times 10^{-11} + 0.05578x_3$
Regency. Tabalong	$\hat{y}_{10} = 0.00024 + 0.18416x_2 + 0.05578x_3 - 0.00059x_4$
Regency. Tanah Bumbu	$\hat{y}_{11} = 7.63458 \times 10^{-5} + 0.05907x_2 + 0.05578x_3 - 1.0887 \times 10^{-4}x_4$
Regency. Tanah Laut	$\hat{y}_{12} = -4.23414 \times 10^{-5} - 0.0368x_2 + 0.05578x_3 + 1.3466 \times 10^{-4}x_4$
Regency. Tapin	$\hat{y}_{13} = 3.66649 \times 10^{-5} + 0.05578x_3$

By the results of the GWR and MGWR, it is obtained which variables significantly affect each region, so which areas can be grouped which are affected by the same factors. These groupings can be displayed as a thematic map in Figure 2.

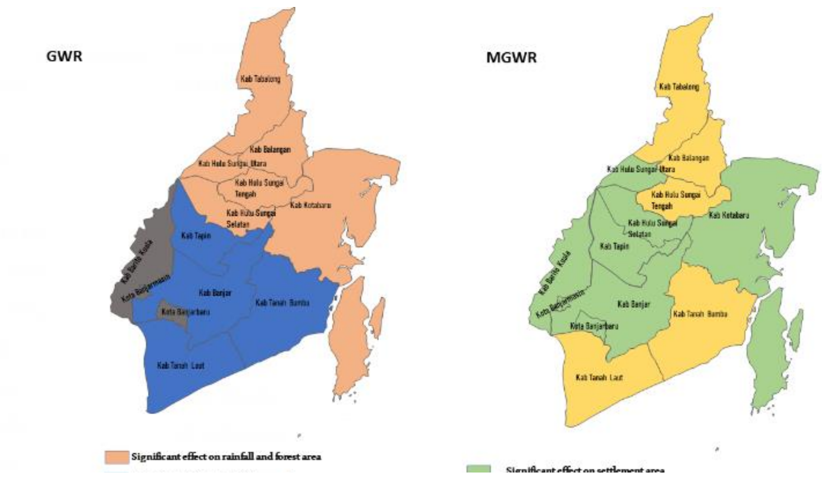

Fig 2. Mapping of the Causes of Floods for the GWR and MGWR Models

Based on Figure 2, the results of variables that significantly influence the occurrence of flooding in South Kalimantan are obtained. In the GWR model, six areas are particularly affected by rainfall (x_2) and forest area (x_4), namely Tabalong Regency, Balangan Regency, Hulu Sungai Selatan Regency, Hulu Sungai Utara Regency, Hulu Sungai Tengah Regency, and Kotabaru Regency. Meanwhile, four regions, namely Tapin Regency, Banjar Regency, Tanah Bumbu Regency, and Tanah Laut Regency, are affected by rainfall (x_2). On the other hand, some areas are unaffected by significant variables, including Barito Kuala Regency, Banjarmasin City, and Banjarbaru City.

The MGWR model differs because the residential size affects the entire area (x_3). In addition, five areas are affected by rainfall (x_2), residential area (x_3) and forest area (x_4), namely Tanah Laut Regency, Tanah Bumbu Regency, Hulu Sungai Tengah Regency, Balangan Regency, and Tabalong Regency.

3.4 Model evaluation

After the flood events in South Kalimantan were modeled using regression, GWR, and MGWR, the results of comparing the three models are presented in Table 10.

Table 10. Evaluation Model

Model	Value R^2
Regression	0.3616
GWR	0.9562
MGWR	0.8857

Based on Table 10, the results of the coefficient of determination for each model will be compared to see the model's performance in explaining flood cases. The GWR model's conclusion coefficient was superior to that of the regression and MGWR methods, which was 0.9562. Because each area is modeled according to the factors that significantly influence the occurrence of floods, on the other hand, the MGWR model has a lower coefficient of determination because it is modeled based on variables that have an overall effect, even though some areas do not experience flooding but are detected to experience flooding.

The results of this study showed that the GWR model was excellent to model flood vulnerability. Some previous studies also resulted in tremendous in the same case, such as research by Lin and Billa [33] regarding the effectiveness of GWR in modeling flood vulnerability, where the model formed can explain 0.91 factors that cause flooding, namely the topographic position index, topographic humidity index, and drainage. Another study by Sifriyani and Ruslan [34] used GWR to analyze the factors causing flooding in Samarinda City. The causes of flooding in Samarinda City include regional topography, population growth, high rainfall and tides that impact the Mahakam River basin. These factors significantly affect the occurrence of floods, which are described through the R^2 value of 92%.

3.5 Conclusion

From the GWR model, it was found that flood disasters are locally caused by rainfall and forest areas. Meanwhile, in the MGWR model, the highest flood cases globally or overall are caused by a residential area, while locally, it is caused by rainfall and forest area. The resulting GWR and MGWR models show that the area prone to flooding is the Balangan Regency. By comparing the results of the modeling evaluation, it can be concluded that in modeling the flood case in South Kalimantan, the GWR model with Adaptive Gaussian weights is superior to the linear regression and MGWR models. This is indicated by the coefficient of determination, which is the largest of the other models at 0.9562.

References

1. D. Krisdiyanto, Apl. J. Apl. Ilmu-ilmu Agama, **20**, 2, 159–181 (2021)
2. R. N. Rachmawati, Procedia Comput. Sci., **179**, 2020, 330–336 (2021)
3. B. E. Hilft, *WeltRisikoBericht 2018*. Germany, 2018.
4. I. F. U. Muzayanah, H. H. Lean, D. Hartono, K. D. Indraswari, and R. Partama, Heliyon, **8**, 9, e10634 (2022)
5. M. Yu, C. Yang, and Y. Li, Geosci., **8**, 5 (2018)
6. A. B. Gertz, J. B. Davies, and S. L. Black, Risk Anal., **39**, 6, 1314–1341 (2019)
7. R. Ramadhan *et al.*, Remote Sens. Appl. Soc. Environ., **28**, 100827 (2022)
8. BPBD, "Rekapitulasi Banjir Menurut Kabupaten/Kota," *Badan Penanggulangan Bencana Daerah*, 2022. https://data.kalselprov.go.id/dataset/data/1042 (accessed Sep. 30, 2022).
9. M. Zulaeha, L. Ariany, A. H. Dwifama, R. A. Falmelia, and M. S. Ridhani, "Mitigasi

Bencana Perspektif Kebijakan Publik Dalam Penanggulangan Benana Banjir di Kalimantan Selatan," *Pros. Semin. Nas. Lingkung. Lahan Basah*, vol. 7, no. 3, pp. 150–159 (2022)
10. P. Prihartini, M. Aini, N. Sya'diah, and R. F. Tazkianida, J. Manaj. Bencana, **7**, 1, 37–44 (2021)
11. A. Priagung, Al Qisthas, vol. **13**, 1, 63–76 (2021)
12. A. A. Noer Dwi, A. Fithria, and K. Kissinger, J. Hutan Trop., **9**, 1, 88, (2021)
13. R. C. Puspitarini, J. Ilmu Sos. dan Polit., **1**, 1, 1–10 (2021)
14. A. Tenri Sompa *et al.*, J. Empower. Community Serv., **1**, 1, 31–36 (2021)
15. H. Yu, H. Gong, B. Chen, K. Liu, and M. Gao, Sci. Total Environ., **738**, 139405 (2020)
16. B. Lu, C. Brunsdon, M. Charlton, and P. Harris, Int. J. Geogr. Inf. Sci., **31**, 5, 982–998 (2017)
17. T. Purwaningsih, C. S. Prajaningrum, and M. Anugrahwati, *Int. J. …*, **2**, 2, 14–27 (2018)
18. X. Chen, M. Emam, L. Zhang, R. Rifhat, L. Zhang, and Y. Zheng, Prev. Med. Reports, **35**, 102362 (2023)
19. C. Wu, G. Liu, and C. Huang, Arch. Agron. Soil Sci., **63**, 7, 928–941 (2017)
20. F. Chen and C. L. Mei, Econ. Model., **94**, 737–747 (2021)
21. C. H. Wei and F. Qi, Econ. Model., **29**, 6, 2615–2620 (2012)
22. C. Zeng *et al.*, Geoderma, **281**, 69–82 (2016)
23. Z. Shabrina, B. Buyuklieva, and M. K. M. Ng, Geogr. Anal., **53**, 4, 686–707 (2021)
24. A. K. Bera and S. G. Kangalli Uyar, J. Eur. Real Estate Res., **12**, 2, 227–249 (2019)
25. L. Chao, K. Zhang, Z. Li, Y. Zhu, J. Wang, and Z. Yu, J. Hydrol., **558**, 275–289 (2018)
26. A. C. Rencher and G. B. Schaalje, *Linear Models in Statistics*, vol. 96, no. 455. New Jersey: John Wiley & Sons, LTD, 2001.
27. H. H. Nuha and A. A. Absa, J. Online Inform., **7**, 1, 1 (2022)
28. D. Kusnandar, N. N. Debataraja, and S. Fitriani, J. Apl. Stat. Komputasi Stat., **13**, 1, 9–16 (2021)
29. S. W. Tyas, Gunardi, and L. A. Puspitasari, MethodsX, **10**, 102002 (2023)
30. H. Wu and J.-T. Zhang, A John Wiley & Sons,Inc, (2006)
31. T. M. Oshan, Z. Li, W. Kang, L. J. Wolf, and A. Stewart Fotheringham, ISPRS Int. J. Geo-Information, **8**, 6 (2019)
32. D. Chicco, M. J. Warrens, and G. Jurman, PeerJ Comput. Sci., **7**, 1–24 (2021)
33. J. M. Lin and L. Billa, Environ. Adv., **6**, 100118 (2021)
34. Sifriyani and Ruslan, AIP Conf. Proc., **2554**, 1, 30002 (2023)