Implementation of Robust Mixed Geographically and Temporally Weighted Regression for Crime Rate Prediction in East Java

Affa Lelira Ibrahim(1), Trimono(2), Aviolla Terza Damaliana(3)

[1,2,3]Sains Data, Fakultas Ilmu Komputer,
Universitas Pembangunan Nasional "Veteran" Jawa Timur

Jl. Rungkut Madya, Surabaya, Jawa Timur
e-mail: 22083010065@student.upnjatim.ac.id(1) , trimono.stat@upnjatim.ac.id(2),
aviolla.terza.sada@upnjatim.ac.id(3)

ABSTRAK

Total kejahatan di Provinsi Jawa Timur mencapai 60,102 kasus pada tahun 2024, yang menunjukkan adanya variasi spasial dan temporal dalam pola kriminalitas. Karakteristik data ini bersifat heterogen dan mengandung beberapa nilai ekstrem (*outlier*) yang dapat memengaruhi stabilitas estimasi model regresi standar. Penelitian ini menerapkan model Robust Mixed Geographically and Temporally Weighted Regression (RMGTWR) berbasis MM-Estimator untuk menganalisis pola kejahatan yang bersifat spasial dan temporal di Jawa Timur selama periode 2020–2024 serta mengevaluasi kemampuan model dalam menangani data yang heterogen dan mengandung pencilan (outlier). Hasil evaluasi menunjukkan model RMGTWR menghasilkan R^2 sebesar 0.84 dan MSE 1912.247. Pendekatan *robust* ini memberikan estimasi parameter yang lebih stabil dan tidak terdistorsi oleh 14 observasi pencilan yang teridentifikasi, dibandingkan model non-robust. Temuan menunjukkan bahwa IPM dan Gini Ratio memiliki pengaruh signifikan secara global, sedangkan variabel Rasio Polisi memiliki pengaruh yang bervariasi secara lokal. Hasil ini mengonfirmasi bahwa pendekatan RMGTWR mampu mengakomodasi heterogenitas data sekaligus memitigasi distorsi akibat pencilan dalam analisis kriminalitas di Jawa Timur.
Kata kunci: Kejahatan; Jawa Timur; RMGTWR; MM-*Estimator*; *Outlier*

ABSTRACT

The total number of crimes in East Java Province reached 60,102 cases in 2024, indicating spatial and temporal variation in crime patterns. These data characteristics are heterogeneous and contain several extreme values (outliers) that may affect the stability of standard regression model estimates. This study applies the Robust Mixed Geographically and Temporally Weighted Regression (RMGTWR) model to examine spatiotemporal crime patterns in East Java during the 2020–2024 period and to evaluate the model's ability to handle outliers in heterogeneous data. The evaluation results show that the RMGTWR model produced an R^2 of 0.84 and an MSE of 1912.247. This robust approach provides more stable parameter estimates and reduces the influence of the 14 identified outlier observations compared with the non-robust model. The findings indicate that the Human Development Index (HDI) and the Gini Ratio have significant global effects, while the Police Ratio variable shows varying local effects. These results confirm that the RMGTWR approach accommodates data heterogeneity while mitigating distortions caused by outliers in crime analysis in East Java.
Keywords: *Crime; East Jawa; RMGTWR; MM-Estimator; Outlier.*

INTRODUCTION

Socio-economic development in Indonesia has created challenges for security stability, reflected in the country's ranking of 52 in the Global Peace Index 2023 [1]. At the regional level, East Java Province recorded the third highest number of crimes in Indonesia in 2024 with a total of 60,102 cases) [2]. This indicates that crime is a complex phenomenon influenced by socio-economic factors and regional heterogeneity, and is closely related to the achievement of SDG 16 on peace, justice, and strong institutions as well as the national development agenda in the Long-Term National Development Plan 2025–2045 [3].

To account for spatial variation in crime rates, the Geographically Weighted Regression (GWR) model is often used because it captures spatial heterogeneity through local parameter estimation [4]. However, since crime patterns also change over time, Geographically and Temporally Weighted Regression (GTWR) was developed to model spatial and temporal variation simultaneously [5]. Furthermore, Mixed Geographically and Temporally Weighted Regression (Mixed GTWR) incorporates both global and local variables in a single model, producing more efficient and stable estimates [6]. Previous studies have demonstrated its effectiveness, such as in analyzing the spread of Hand, Foot, and Mouth Disease in Mongolia, where it achieved higher accuracy than the standard GTWR model [7].

However, both GTWR and Mixed GTWR models remain sensitive to outliers in spatiotemporal data. Extreme values can affect parameter estimation, leading to less stable and potentially biased models. Similar issues have motivated the development of robust approaches for GTWR, such as the Robust Improved GTWR based on the M-Estimator [8]. Nevertheless, this approach has limitations, as the M-Estimator alone may not be sufficient to handle a high proportion of outliers.

Based on this context, the study addresses a research gap by applying the Robust Mixed Geographically and Temporally Weighted Regression (RMGTWR) using the MM-Estimator, which combines the S-Estimator to obtain a robust scale with the M-Estimator to improve estimation efficiency. This approach has a high breakdown point and produces more stable estimates against the influence of extreme data. The model is applied to predict crime rates in East Java, considering socio-economic and law enforcement factors, and is expected to provide more robust predictions that support spatiotemporal-based preventive policies.

METHOD

Data and Data Sources

This study uses secondary data from official publications of the East Java Provincial Statistics Agency, covering 38 regencies/cities during 2020–2024. The dependent variable is the crime rate, measured as crimes per 100,000 population to control for population differences across regions. Spatial coordinates use the Universal Transverse Mercator (UTM) system to ensure accurate distance calculations. The variables used in this study are as follows:

Table 1. Research Variables

Symbol	Variable	Type of Data
Y	Crime Rate per 100,000 Population	Ratio
X_1	Percentage of Poor Population (PPM)	Ratio (%)

X_2	Open Unemployment Rate (TPT)	Ratio (%)
X_3	Human Development Index (IPM)	Index (0-100)
X_4	Police Ratio (RPO)	Ratio
X_5	Gini Ratio (GR)	Ratio (0-1)
u_i	UTM Easting coordinate	Meter
v_i	UTM Northing coordinate	Meter

Analysis Steps

The study follows these steps:

1. Collect data from the East Java Statistics Office, combine into a dataset of 190 observations, and perform descriptive and spatial exploratory analyses to understand initial data characteristics.
2. Test for multicollinearity using the Variance Inflation Factor (VIF) to detect high correlations among independent variables [9], calculated as $VIF_j = \frac{1}{1-R_j^2}$. The null hypothesis assumes no multicollinearity and is rejected if $VIF_j > 10$, indicating multicollinearity.
3. Test spatial dependency (Moran's I) to identify spatial correlation in the residuals of the global model [10], calculated as $I = \frac{n}{\mathbf{W}} \cdot \frac{\sum_{i=1}^{n}\sum_{j=1}^{n} w_{ij} e_i e_j}{\sum_{i=1}^{n} e_i^2}$ where w_{ij} is the spatial weight between locations i and j, $\mathbf{W}$ is spatial weight, and e_i is the residual at location i. The null hypothesis assumes no spatial dependency and is rejected if $|Z(I)| > Z_{\alpha/2}$, or $p - value <$ 0,05, where $Z(I) = \frac{I - E[I]}{\sqrt{Var(I)}}$.
4. Test spatial heterogeneity using the spatially weighted Breusch–Pagan (BP) test, which incorporates the spatial weight matrix (W) to detect spatial heteroskedasticity in the residuals across regions [9], with the test statistic BP $= \frac{1}{2}\boldsymbol{f}^T\mathbf{Z}(\mathbf{Z}^T\mathbf{Z})^{-1}\mathbf{Z}^T\boldsymbol{f} + \frac{1}{T}\left(\frac{e^T\mathbf{W}e}{\sigma^2}\right)$ where f is the vector of squared residuals, $\mathbf{Z}$ is the matrix of explanatory variables, $\mathbf{W}$ is the spatial weight matrix, e is the residual vector, and σ^2 is the residual variance. The null hypothesis assumes homogeneous residual variance and is rejected if $BP > \chi^2_{\alpha,p-1}$, or $p - value < 0{,}05$.
5. Conducting a temporal heterogeneity test by examining the interaction between explanatory variables and the time factor using model comparison via ANOVA (partial F-test).
6. Determine global and local variables using residual bootstrap tests.
7. Identify outliers using R-Student in the Mixed GTWR model as a basis for applying the robust approach. The test statistic is formulated as $t_i = \frac{e_i}{s_i\sqrt{1-h_{ii}}}$ where e_i is the residual of observation i, s_i is the estimated standard error of observation i, and h_{ii} is the leverage value from the hat matrix. The null hypothesis assumes no outliers and is rejected if $|t_i| > t_{\frac{\alpha}{2},n-p-1}$, or $p - value < 0{,}05$, indicating the presence of outliers [11].
8. Estimate the Robust Mixed GTWR model with the MM-estimator then obtain the final parameters $\hat{\beta}$ via Weighted Least Squares (WLS) using combined weights $\mathbf{W}_{i,total}$.

9. Test the significance of global parameters in the Robust Mixed GTWR model to assess the effect of independent variables that are constant [12], with the test statistic $t_g = \frac{\hat{\beta}_g}{\hat{\sigma}_g\sqrt{g_{kk}}}$ where $\hat{\beta}_g$ is the estimated global coefficient, $\hat{\sigma}_g$ is its standard error, and g_{kk} is the diagonal element of the covariance matrix of global parameters. The null hypothesis assumes the global parameter is not significant and is rejected if $T_{hit} > t_{\alpha/2,df}$, dengan $df = n - 2 \cdot tr(\mathbf{L}_M) + tr(\mathbf{L}_M^T\mathbf{L}_M)$
10. Test the significance of local parameters in the Robust Mixed GTWR model to assess the effect of variables that vary spatiotemporally [12], with the test statistic $t_l = \frac{\hat{\beta}_l(u_i,v_i,t_i)}{\hat{\sigma}\sqrt{m_{kk}}}$ where $\hat{\beta}_l(u_i, v_i, t_i)$ is the estimated local coefficient at spatial coordinates (u_i, v_i) and time t_i, and m_{kk} is the diagonal element of the covariance matrix of local parameters. The null hypothesis assumes the local parameter at location i is not significant and is rejected if $T_{hit} > t_{\alpha/2,df}$, with $df = n - 2 \cdot tr(\mathbf{L}_M) + tr(\mathbf{L}_M^T\mathbf{L}_M)$
11. Evaluate the Robust Mixed GTWR model using MSE and R^2, where MSE is calculated as $MSE = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2$ and R^2 is calculated as $R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i-\hat{y}_i)^2}{\sum_{i=1}^{n}(y_i-\bar{y})^2}$. The results are then compared with those obtained from the GTWR and Mixed GTWR models.

Determination of Global and Local Variables Using Bootstrap Test

A residual-based bootstrap test is used to determine whether a variable is global or local in the Mixed GTWR model. The null hypothesis assumes the tested variable is global (Mixed GTWR), while the alternative hypothesis assumes it is local (GTWR) [7]. The predicted values for the global and local models are expressed as $\hat{\mathbf{Y}} = \mathbf{L}(h_s, h_t)\mathbf{Y}$ and $\hat{\mathbf{Y}}_M = \mathbf{L}_M(h_s, h_t)\mathbf{Y}$ respectively, where Y is the vector of observed response values, $\hat{Y}$ and $\hat{Y}_M$ denote the predicted values from the GTWR and Mixed GTWR models, $L(h_s, h_t)$ and $L_m(h_s, h_t)$ represent the smoothing matrices for the GTWR and Mixed GTWR models. Furthermore, the Residual Sum of Squares (RSS) for each model is calculated as shown in Equations (1) and (2).

$$RSS_M(h_s, h_t) = \mathbf{Y}^T\left(\mathbf{I} - \mathbf{L}_M(h_s, h_t)\right)^T\left(\mathbf{I} - \mathbf{L}_M(h_s, h_t)\right)\mathbf{Y} \tag{1}$$

$$RSS = \mathbf{Y}^T\left(\mathbf{I} - \mathbf{L}(h_s, h_t)\right)^T\left(\mathbf{I} - \mathbf{L}(h_s, h_t)\right)\mathbf{Y} \tag{2}$$

where $RSS_M(h_s, h_t)$ and $RSS(h_s, h_t)$ represent the residual sum of squares of the Mixed GTWR and GTWR models, respectively, $\mathbf{I}$ is the identity matrix, and Y^T denotes the transpose of the response vector Y. The test statistic is formulated as shown in Equation (3):

$$Z = \frac{RSS_{H_0} - RSS_{H_1}}{RSS_{H_1}} = \frac{RSS_M(h_s, h_t) - RSS(h_s, h_t)}{RSS(h_s, h_t)} \tag{3}$$

where RSS_{H_0} and RSS_{H_1} denote the residual sum of squares under the null hypothesis (Mixed GTWR) and the alternative hypothesis (GTWR), respectively. Bootstrap data is generated based on the model H_0 with $\mathbf{Y}^* = \mathbf{L}_M(h_s, h_t)\mathbf{Y} + \boldsymbol{\varepsilon}^*$, where Y^* represents the bootstrap response vector and ε^* denotes the bootstrap residuals generated through random resampling. The bootstrap process is repeated B times to obtain the empirical distribution Z^*, with p-value calculated using Eq. (4):

$$p - value = \frac{1}{B}\sum_{i=1}^{B} I(Z_i^* \geq Z) \quad (4)$$

where Z_i^* denotes the bootstrap test statistic at the i-th replication and $I(\cdot)$ is an indicator function that takes the value 1 if the condition inside the parentheses is satisfied and 0 otherwise. The null hypothesis is rejected if p-value < 0.05, in which case the variable is considered local otherwise, it is categorized as global.

Robust Mixed Geographically and Temporally Weighted Regression

Robust Mixed Geographically and Temporally Weighted Regression (Robust Mixed GTWR) extends Mixed GTWR by incorporating a robust regression approach to handle outliers [12]. In this model, global coefficients remain constant, while local coefficients vary across space and time. The Mixed GTWR model is expressed as shown in Equation (5):

$$y_i = \beta_0(u_i, v_i, t_i) + \sum_{i=1}^{p} \beta_g x i_g + \sum_{i=p+1}^{q} \beta_l(u_i, v_i, t_i) x_{il} + \varepsilon_i ; i = 1,2,3, \ldots n \quad (5)$$

where $\beta_g x_{ig}$ are fixed global coefficients, $\beta_l(u_i, v_i, t_i) x_{il}$ are local coefficients that vary across space and time, and ε_i is the error for observation i. In this study, spatiotemporal weights (w_{ij}) are calculated using an Adaptive Gaussian kernel with a fixed number of nearest neighbors per location. The Gaussian kernel ensures weights decrease smoothly with increasing spatial and temporal distance [13]. Mathematically, the weights are expressed in Equation (6):

$$w_{ij}(h_s, h_t) = \exp\left(-{}^{1}/_{2}\left({}^{d_{ij}^S}/_{h_{s(i)}}\right)^2 - {}^{1}/_{2}\left({}^{d_{ij}^T}/_{h_{t(i)}}\right)^2\right), j = 1,2, \ldots, n \quad (6)$$

where d_{ij}^S is the spatial distance between locations, d_{ij}^T is the temporal distance between time points, and $h_{s(i)}, h_{t(i)}$ are the spatial dan temporal bandwidths at location i. The optimal bandwidth was determined using the Corrected Akaike Information Criterion (AICc). Parameter estimation in Mixed GTWR is performed using Weighted Least Squares (WLS), where local coefficients are computed after controlling for the already estimated global components, as formulated using Equation (7):

$$\hat{\beta}_l(u_i, v_i, t_i) = \left(\mathbf{X}_l^T \mathbf{W}_i(h_s, h_t) \mathbf{X}_l\right)^{-1} \mathbf{X}_l^T \mathbf{W}_i(h_s, h_t)(\mathbf{Y} - \mathbf{X}_g \hat{\beta}_g) \quad (7)$$

Here $\mathbf{Y} - \mathbf{X}_g \hat{\beta}_g$ represents the response after removing the global component, so that the estimation of local coefficients captures only spatiotemporal variation. To enhance robustness against outliers, the Robust Mixed GTWR model integrates the MM-Estimator, which combines the robustness of the S-Estimator with the efficiency of the M-Estimator [14]. In general, the MM-Estimator minimizes a loss function based on standardized residuals using Equation (8):

$$min \sum_{i=1}^{n} \rho(u_i), \quad ui = {}^{e_i}/_{\hat{\sigma}_s} \quad (8)$$

The loss function used is the Huber function, defined as $\rho(u) = \frac{1}{2}u^2$ for $|u| \leq c$ and $\rho(u) = c(|u| - \frac{1}{2}c)$ for $|u| > c$, where $c = 1.345$ is a tuning constant used to limit the influence of extreme residuals. From this function, robust weights are obtained and multiplied by the spatiotemporal kernel weights to produce the total weights $\mathbf{W}_{i,total}$. These total weights are then applied in the Weighted Least Squares (WLS) estimation, similar to the Mixed GTWR approach.

The local parameter estimates are calculated using Equation (7). The estimation process is performed iteratively until the parameter $\hat{\beta}$ estimates converge, producing stable results that are robust to the influence of outliers.

RESULT AND DISCUSSION

Descriptive Statistics

Table 2 presents a descriptive overview of crime patterns and characteristics in East Java Province.

Table 2. Descriptive Statistics

Variable	Mean	Std. Dev	Min	Max
Y	133.26	113.19	7.30	621.06
X_1	10.55	4.41	3.06	23.76
X_2	5.02	1.79	1.36	10.97
X_3	74.05	4.75	64.75	84.69
X_4	115.15	94.23	19.45	434.96
X_5	0.34	0.03	0.23	0.44

The descriptive statistics show that crime rates (Y) have a mean of 133.26 with high variation (7.30–621.06), indicating heterogeneity and potential outliers. The Human Development Index (IPM, X_3) has a mean of 74.05, reflecting differences in development across regions, while the police ratio (X_4) shows very high variation (maximum 434.96), and the Gini ratio (X_5) has a mean of 0.34 (range 0.23–0.44). Overall, the substantial variation in several variables highlights spatiotemporal heterogeneity and the presence of potential outliers.

Spatial Exploration

Spatial exploration was conducted using thematic maps to illustrate the distribution patterns of crime rates in each regency/city for each year, as shown in Figure 1.

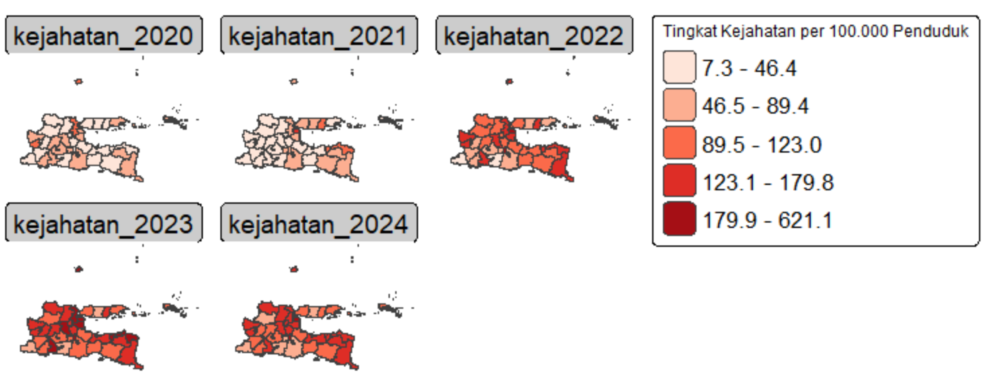

Figure 1. Crime Rate Distribution per 100,000 Population

Thematic map exploration shows variation in crime rates across regencies/cities in East Java during 2020–2024. Increased intensity in certain areas indicates that crime is not evenly distributed but concentrated in specific regions, reflecting spatial heterogeneity and temporal dynamics.

Multicollinearity Test

The multicollinearity test results show that all independent variables have Variance Inflation Factor (VIF) values below 10, namely 3.16 for X_1, 1.47 for X_2, 4.01 for X_3, 1.58 for X_4, and 1.61 for X_5. These results indicate that no serious multicollinearity problem is present among the explanatory variables.

Spatial Dependence Test (Moran's I)

The Moran's I test produced a value of 0.2669 with a standard deviate of 10.34 and a p-value $< 2.2 \times 10^{-16}$. This positive and significant result indicates spatial autocorrelation, meaning regions with high (or low) crime rates tend to cluster with areas showing similar characteristics.

Spatial and Temporal Heterogeneity Tests

The spatially weighted Breusch–Pagan test produced BP 187.41 with p-value 9.04×10^{-38}, indicating spatial heteroskedasticity and violation of the homogeneous variance assumption. Temporal heterogeneity was tested using an interaction model between explanatory variables and year, where ANOVA produced F 6.839 with p-value 9.044×10^{-15}, indicating coefficients vary significantly across years. These results confirm spatial and temporal heterogeneity, supporting the use of a spatiotemporal model such as Robust Mixed GTWR.

Determination of Global and Local Variables Using Bootstrap Test

Global and local variables were identified using a bootstrap test with 1000 replications (Table 3).

Table 3. Global Coefficient Test Results Based on 1000 Bootstrap Replications

Null Hypothesis	Related Variable	Observed Z	*P - Value*	Decision
$\beta_0(u,v,t) = \beta_0$	*Intercept*	1.942	0.000	Local
$\beta_1(u,v,t) = \beta_1$	PPM	-0.246	0.643	Global
$\beta_2(u,v,t) = \beta_2$	TPT	-0.022	0.175	Global
$\beta_3(u,v,t) = \beta_3$	IPM	0.639	0.991	Global
$\beta_4(u,v,t) = \beta_4$	RPO	0.700	0.000	Local
$\beta_5(u,v,t) = \beta_5$	GR	0.263	0.138	Global

The bootstrap results in Table 3 indicate that the intercept and the police ratio (RPO) are local variables ($p < 0.05$), meaning their coefficients vary across space and time. This suggests differences in baseline crime levels and context-dependent effects of police presence across districts and cities. In contrast, PPM, TPT, IPM, and GR are identified as global variables, indicating relatively consistent effects across the study area. Next, the Mixed GTWR model is estimated to detect outliers through diagnostic analysis. This step aims to identify deviant observations that may affect parameter stability, forming the basis for applying the Robust Mixed GTWR model to increase robustness against outliers.

Outlier Detection (R-Student)

A summary of the detected observations is presented in Table 4.

Table 4. Outlier Detection Results in the Mixed GTWR Model

No	Regency/City	Year	*R-Student*
1	Pasuruan City	2023	4.98
2	Probolinggo City	2023	3.67
⋮	⋮	⋮	⋮
13	Mojokerto City	2020	-2.21
14	Blitar City	2024	2.007

The R-Student test identified 14 out of 190 observations with absolute values greater than 2, indicating potential outliers. Because R-Student accounts for leverage, they may influence parameter estimation. Therefore, the Robust Mixed GTWR model with MM-estimator was applied to reduce the effect of extreme observations without removing them.

Robust Mixed Geographically and Temporally Weighted Regression

The Robust Mixed GTWR modeling began with determining the optimal Adaptive Gaussian kernel bandwidth based on the lowest AICc, resulting in a spatial bandwidth h_s of 11 nearest neighbors and a temporal bandwidth h_t of 1 year, with an AICc value of 2065.881. The robust scale from S-estimation, 32.7556, was used as a fixed scale in the M-estimation. Parameter estimation via IRLS reached convergence at 0.000007 on the 20th iteration, indicating stable and robust results against spatiotemporal heterogeneity and outliers. The coefficients were then classified into global and local parameters (Table 5 and 6).

Table 5. Global Estimated Parameter

Variable	Parameter	Estimation Value
PPM	$\hat{\beta}_1$	-0.065
TPT	$\hat{\beta}_2$	-3.28
IPM	$\hat{\beta}_3$	6.14
GR	$\hat{\beta}_5$	364.34

Based on Table 5, PPM and TPT have negative global coefficients, indicating that increases in these variables tend to reduce crime rates overall. In contrast, IPM and GR have a positive influence on crime levels.

Table 6. Summary of Local Estimated Parameter

Variable	Parameter	Min	Max	Mean	Std
Intercept	$\hat{\beta}_0$	-544.35	-421.27	-497.29	23.482
RPO	$\hat{\beta}_4$	0.0203	1.0567	0.578	0.2413

Table 6 shows that the local parameters $\hat{\beta}_0$ and $\hat{\beta}_4$ exhibit considerable variation across regions and time. This is evident from the minimum–maximum range and relatively high standard deviations, indicating spatiotemporal heterogeneity and dynamics in these parameters. For example, in Surabaya City and Sumenep Regency in 2024, the model is as shown in Equations (9) and (10):

$$\hat{Y}_{Surabaya\_2024} = -476.12 - 0.065X_1 - 3.28X_2 + 6.14X_3 + 1.09X_4 + 364.34X_5 \quad (9)$$

$$\hat{Y}_{Sumenep\_2024} = -476.47 - 0.065X_1 - 3.28X_2 + 6.14X_3 + 0.93X_4 + 364.34X_5 \quad (10)$$

In 2024, the police ratio (X_4) coefficient was 1.09 in Surabaya and 0.93 in Sumenep, indicating a positive association between the police ratio and crime rates. However, this relationship may reflect reactive policing, where higher police presence is deployed in response to elevated crime levels rather than representing a direct causal effect. The difference in coefficient values between the two regions also suggests spatial heterogeneity, potentially related to variations in social characteristics, population density, and institutional capacity.

Global Significance Test Robust Mixed GTWR

The results of the global significance test are presented in Table 7.

Table 7. Global Significance Test Results

Variable	t-Statistic	P-value	Conclusion
PPM (X_1)	-0.028	0.977	Not Significant
TPT (X_2)	-0.765	0.44	Not Significant
IPM (X_3)	2.73	0.006	Significant
GR (X_5)	1.98	0.048	Significant

The global significance test shows that IPM and GR significantly influence crime rates, while PPM and TPT are not significant. The positive IPM coefficient suggests that more developed regions tend to experience higher crime, possibly due to increased economic activity and criminal opportunities. Meanwhile, the positive effect of GR supports the relative deprivation perspective [15], indicating that inequality may contribute to higher crime. These findings highlight the importance of promoting both development and equitable distribution.

Local Significance Test Robust Mixed GTWR

The results of the local significance test are presented in the map shown in Figure 3.

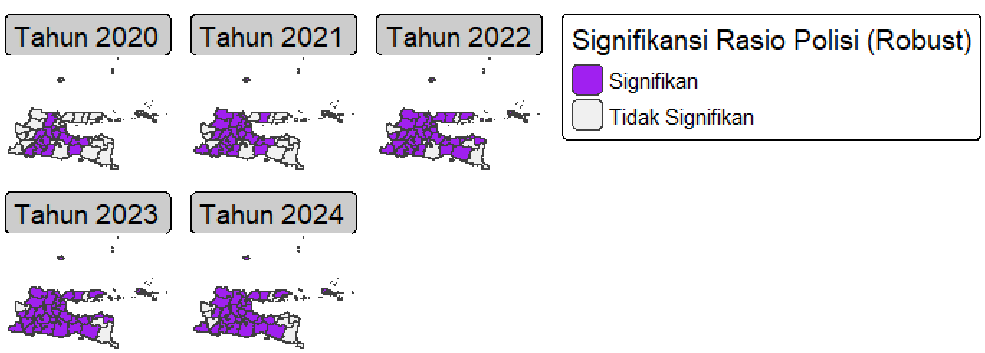

Figure 3. Spatial-Temporal Significance Map of the Police Ratio Variable

Figure 3 shows that in several regions, especially in 2023–2024, the police ratio is positively associated with crime rates. This pattern likely reflects reverse causality, where more police are assigned to high-crime areas. In urban regions, higher police presence and reporting rates may also contribute to this positive correlation, indicating that the relationship is context-dependent.

Model Evaluation

Model performance was evaluated based on R^2 and MSE, with results shown in Table 8.

Table 8. Robust Mixed GTWR Model Evaluation

Model	R^2	MSE
GTWR	0.88	1452.8
Mixed GTWR	0.88	1483.248
Robust Mixed GTWR	0.84	1912.247

The Robust Mixed GTWR model explains 84% of the variation in crime rates. Although its R^2 and MSE are slightly lower than those of GTWR and Mixed GTWR, the robust approach reduces the influence of outliers and produces more stable estimates. It also changes the statistical inference for the Gini variable, indicating that inequality effects are sensitive to extreme observations. Overall, the robust model provides more reliable inference despite a slightly higher prediction error.

CONCLUSION

This study shows that the Robust Mixed GTWR model with MM-Estimator can address spatiotemporal heterogeneity and outliers in crime data in East Java during 2020–2024. By combining S-estimation and M-estimation, the model produced stable parameter estimates (converging at the 20th iteration) with an R^2 of 0.84, indicating strong explanatory capability despite the presence of 14 outliers. The results reveal that IPM and the Gini Ratio (GR) have consistent global effects, while the Police Ratio (RPO) and the intercept vary locally across regions and time, highlighting the contextual nature of crime dynamics. These findings suggest that crime prevention policies should consider spatial–temporal differences across regions. Law enforcement allocation may need to prioritize areas with higher crime risk, particularly urban regions where socio-economic development and inequality coexist. This study is limited by the use of aggregated regional data, which may not capture unreported crime. Future research could incorporate more detailed socio-economic variables and explore alternative weighting schemes to improve model performance in dense urban areas.

REFERENCE

[1] Institute for Economics & Peace. Global Peace Index 2024: Measuring Peace in a Complex World, Sydney, June 2024. Available from: http://visionofhumanity.org/resources

[2] Badan Pusat Statistik, Statistik Kriminal 2024, Jakarta: BPS RI, 2024. [Online]. Available: https://www.bps.go.id.

[3] Sustainable development goal 16: Perdamaian, Keadilan Dan Kelembagaan Yang Tangguh perserikatan Bangsa-Bangsa di Indonesia," https://indonesia.un.org/id/sdgs/16/key-functions.

[4] H. Leitner and E. Sheppard, "Theorizing Globally, But Analyzing Locally: The Importance Of Geographically Weighted Regression In Crime Analysis", Crime Science, vol. 11, no. 1, pp. 1–14, 2022.

[5] X. X. Zhu *et* al, "Geographically and Temporally Weighted Regression in Assessing Dengue Fever Spread Factors in Yunnan Border Regions," *Biomedical and Environmental Sciences*, vol. 37, no. 5, pp. 511–520, May 2024, doi: 10.3967/bes2024.056.

[6] J. Liu *et al.*, "A mixed geographically and temporally weighted regression: Exploring spatial-temporal variations from global and local perspectives," *Entropy*, vol. 19, no. 2, 2017, doi: 10.3390/e19020053.

[7] Z. Hong, C. Mei, H. Wang, and W. Du, "Spatiotemporal effects of climate factors on childhood hand, foot, and mouth disease: a case study using mixed geographically and temporally weighted regression models," *International Journal of Geographical Information Science*, vol. 35, 2021, doi: 10.1080/13658816.2021.1882681.

[8] S. A. Ania, N. Ilyas, and E. T. Herdiani, "Robust Spatial Temporal Analysis with Improved Geographically and Temporally Weighted Regression Model of Dengue Incidence Rate," *Communications in Mathematical Biology and Neuroscience*, 2024, doi: 10.28919/cmbn/8376.

[9] N. Nessyana Debataraja, D. Kusnandar, R. Mahalalita, and N. Imro'ah, "Penerapan Model Geographically and Temporally Weighted Regression Pada Kecelakaan Lalu Lintas," *Jurnal Siger Matematika*, vol. 02, pp. 19–24, 2021.

[10] Y. Rahkmawati, R. Faulina, and I. Nur Laily Fitriana, "Spatial Analysis of Gross Enrollment Ratio in Higher Education Using the Geographically Weighted Regression Approach," *J Statistika*, vol. 18, no. 2, 2025.

[11] M. Jainudin, M. N. Hayati, and D. I. Purnamasari, "Deteksi Pencilan Spasial pada Data Kandungan Klorida di Sungai Mahakam Wilayah Samarinda Kalimantan Timur," *Jurnal EKSPONENSIAL*, vol. 10, no. 2, 2019.

[12] A. Hujjah Asianingrum and A. Djuraidah, "Robust Mixed Geographically and Temporally Weighted Regression to Modeling the Percentage of Poverty Population in Java in 2012-2018," *International Journal of Sciences: Basic and Applied Research (IJSBAR)*, vol. 53, no. 2, pp. 186–197, 2020.

[13] G. Khusnul Ma'rifah, M. Idhom, and Trimono, "Implementation of Geographically and Temporally Weighted Regression with Cross Validation and Generalized Cross Validation Methods for Deforestation Modeling in Kalimantan," *J Statistika*, vol. 18, no. 1, 2025

[14] A. Rahmah Putri, H. Yozza, and F. Yanuar, "Regresi Robust MM-Estimator untuk Memodelkan Jumlah Kematian Balita di Provinsi Jawa Timur Tahun 2017," *Jurnal Matematika UNAND*, vol. 10, no. 1, pp. 71–78, 2021.

[15] E. Itskovich, "Economic Inequality, Relative Deprivation, and Crime: An Individual-Level Examination," *Justice Quarterly*, vol. 42, no. 4, pp. 637–658, 2025, doi: 10.1080/07418825.2024.2435859.