Aksioma: Jurnal Matematika dan Pendidikan Matematika
Volume 17, No. 1, April 2026, pp. 35-50
ISSN: 2579-7646 (e) | 2086-2725 (p)
https://doi.org/10.26877/3v0x5p12

Modeling the frequency of traffic accident compensation claims by jasa raharja using the GSTARX model

Auliya Hasna Rahadatul 'Aisy[1], Swasono Rahardjo[2]

[1,2]Mathematics Department, Universitas Negeri Malang, Malang, Jawa Timur, Indonesia
Correspondence: swasono.rahardjo.fmipa@um.ac.id

Received: March 10, 2026 | **Revised**: Apr 14, 2026 | **Accepted**: Apr 17, 2026 | **Published Online**: Apr 28, 2026

Abstract

Fluctuations in traffic accident frequency increase PT Jasa Raharja's exposure to claims and financial risks, while total compensation claims are stochastic because they are influenced by incident frequency and accident severity. Although the Generalized Space-Time Autoregressive with Exogenous Variable (GSTARX) model has been widely used to model spatial-temporal data, its application remains primarily in macroeconomics, simulation studies, and non-insurance phenomena. Hence, its application to model the frequency of traffic accident compensation claims, influenced by spatio-temporal relationships and external factors such as calendar variations, remains relatively limited. Therefore, this study aims to model and forecast the frequency of traffic accident compensation claims in four police districts, namely Batu Police District, Malang City Police District, Malang Police District, and Pasuruan Police District, for the period 2020-2024 using the GSTARX model. The data used are monthly secondary data from PT Jasa Raharja Malang Branch, with exogenous variables in the form of calendar variables. The analysis was conducted through data characteristic identification, stationarity testing, spatial-temporal correlation analysis using the Cross Correlation Function (CCF), model identification, location-weight determination, parameter estimation using Seemingly Unrelated Regression (SUR), and forecasting and model accuracy evaluation. The results of the study show that the GSTARX(2,1)(0,0,0) model with CCF(0)-based neighborhood location weights and cross-correlation normalization location weights produces relatively identical forecasting performance.

Keywords: GSTARX; Location Weighting; Traffic Accident Claims; Forecasting

How to Cite: Aisy, A.H.R., Rahardjo, S. (2026). Modeling the frequency of traffic accident compensation claims by jasa raharja using the GSTARX model. *Aksioma: Jurnal Matematika dan Pendidikan Matematika, 17(1)*, 35-50. https://doi.org/10.26877/3v0x5p12

INTRODUCTION

Traffic accidents are among the inherent risks of land transportation innovation, causing fatalities, injuries, and financial losses to affected parties (Go & Mauliddin, 2024). The Law of the Republic of Indonesia Number 22 of 2009 on Road Traffic and Transportation defines a traffic accident as an unexpected and unintentional event on the road involving a vehicle with or without other road users, resulting in human casualties and/or property damage. Based on data from the Central Statistics Agency (BPS) sourced from the Indonesian National Police, the frequency of traffic accidents during the 2020-2024 period fluctuated, with an increasing trend, as reflected in the annual number of accidents exceeding 150,000 in the last two years (BPS, 2025). These events indicate that an increase in traffic accident frequency implies a greater risk of claims for insurance companies (Aulia et al., 2025).

As a form of state protection for traffic accident victims, the government established PT Jasa Raharja, a State-Owned Enterprise (BUMN), as a social insurance institution responsible for providing compensation to victims. This protection is provided through two insurance programs: the Passenger Accident Insurance for Public Transportation, regulated by Law Number 33 of 1964, and the Legal Liability Insurance for Third Parties, regulated by Law Number 34 of 1964. The amount of compensation has been stipulated in the Minister of Finance Regulation Numbers 15/PMK.010/2017 and 16/PMK.010/2017; however, the total compensation claims to be paid are stochastic because they are influenced by the frequency of incidents and the severity of traffic accidents (PT Jasa Raharja, 2024). Therefore, optimal risk management is needed to support the planning of claim compensation reserve funds, including through claim frequency modeling.

Aulia et al. (2025) demonstrated that the accuracy of claim frequency modeling affects premium setting, technical reserve planning, and risk management in insurance companies. In their study, claim frequency was modeled using Poisson regression to estimate the number of life insurance claims, accounting for the number of active policies managed by the company. Nevertheless, this approach remains non-spatial and does not account for inter-regional relationships or temporal dynamics, so it does not fully capture the characteristics of claim frequency.

The Space Time Autoregressive (STAR) model developed by Pfeifer and Deutsch (1980) is one approach for modeling data with interregional and temporal dependencies. This model combines interregional and temporal interdependencies under the assumption that all observed regions have the same characteristics (homogeneity), so that the autoregressive and spatial parameters apply equally across regions. However, this homogeneity assumption causes model limitations when applied to data with different regional characteristics (heterogeneous). Based on these limitations, the STAR model was subsequently developed into the Generalized Space Time Autoregressive (GSTAR) model by Borovkova et al. (2002), which assumes that all observed regions have heterogeneous characteristics and a constant error variance (Suryani & Saputro, 2018; Gustiasih & Saputro, 2018).

In addition to being influenced by interregional and temporal relationships, spatio-temporal time series data can also be influenced by external factors or exogenous variables. The development of the GSTAR model by incorporating exogenous variables is known as the Generalized Space-Time Autoregressive with Exogenous Variables (GSTARX) model (Fadlurohman et al., 2020). This model has been widely applied to model data with spatio-temporal relationships across various fields. Fadlurohman et al. (2020) applied the GSTARX model to inflation in six Cost of Living Survey (SBH) cities in Central Java, using exogenous variables based on calendar variations of Eid al-Fitr, and the results showed that the model captured spatio-temporal inter-regional relationships. Mubarak et al. (2023) used the GSTARX model to model money inflow and outflow at four Regional Representative Offices of Bank Indonesia in East Java, with exogenous variables in the form of Eid al-Fitr and the Consumer Price Index (CPI), and compared four types of location weights, namely uniform weights, inverse distance, cross-correlation normalization, and partial cross-correlation inference normalization. The results showed that the GSTARX model with cross-correlation-normalized weights achieved the best forecasting accuracy. Other research showed that adding calendar variation as an exogenous variable improved forecasting accuracy for spatio-temporal data containing trends and seasonality (Hikmawati et al., 2021). Furthermore, Suhartono et al. (2018) explained that major holiday events, such as Eid al-Fitr, Galungan, and Nyepi, are more appropriately modeled as calendar variation dummy variables in the GSTARX model. In the parameter estimation process, the Seemingly Unrelated Regression (SUR) approach in the GSTARX model can improve estimation efficiency when residuals across regions are correlated (Pramoedyo et al., 2020).

Although the GSTARX model has been widely used to model spatio-temporal data, most of its applications remain focused on macroeconomic data, simulation studies, and non-insurance phenomena. The application of the GSTARX model to the frequency of spatio-temporally structured traffic accident compensation claims remains relatively limited. In fact, the characteristics of compensation claim frequency are influenced not only by inter-regional and temporal relationships but also by external factors such as calendar variations, requiring a modeling approach that can simultaneously accommodate spatio-temporal relationships while accounting for the influence of exogenous variables. In addition, empirical studies on the influence of calendar variations in modeling compensation claim frequency at the Police District (Polres) level are also still limited. Therefore, this study aims to model and forecast the frequency of traffic accident compensation claims using the GSTARX model, treating calendar variations as exogenous variables, and to compare two location-weight approaches: CCF(0)-based neighborhood location weights and cross-correlation normalization location weights.

METHODS

This research uses secondary data in the form of the frequency of traffic accident compensation claims covering four police districts, namely Batu Police District, Malang City Police District, Malang Police District, and Pasuruan Police District, for the period January 2020 to March 2025, obtained from PT Jasa Raharja Malang Branch. Data for the 2020-2024 period is used as in-sample data for model identification, parameter estimation, and diagnostic testing, while data for January to March 2025 is used as out-of-sample data to evaluate the accuracy of the GSTARX model's forecasting results. The endogenous variable is the frequency of traffic accident compensation claims in region i and time t, denoted as $Z_i(t)$, while the exogenous variable are calendar variation indicator variables expressed as dummy variable, namely a dummy for the Eid al-Fitr holiday period and one month after it $\left(D_1(t)\right)$ and a dummy for the Christmas and New Year period, denoted as $\left(D_2(t)\right)$.

The analysis process for developing the GSTARX model for traffic accident compensation claim frequency data was conducted in stages. The analysis began with data characteristic identification using descriptive statistics and time series plots, and with the analysis of interregional relationships using Pearson correlation coefficients (Sugiyono, 2007). Data stationarity was tested for the variance and mean using the Box-Cox transformation, first-order differencing, and the Augmented Dickey-Fuller (ADF) test (Makridakis et al., 1998; Wei, 2006). Spatio-temporal relationships between regions were analyzed using the Cross Correlation Function (CCF), which can be written as (Wei, 2006):

$$\rho_{ij}(k) = \frac{Cov\left(Z_i(t), Z_j(t+k)\right)}{\sigma_i \sigma_j}.$$

The identification of the GSTARX model was conducted in two stages: the identification of the GSTAR model and the identification of exogenous variables. The

GSTAR model was identified by determining the time order based on the smallest Akaike Information Criterion (AIC) value of the VAR model (Zhao et al., 2022), as well as setting the spatial order (λ), which is generally limited to ordering one $(\lambda = 1)$ because higher orders tend to be difficult to interpret (Rahmani et al., 2025); thus, in this study, a spatial order of $\lambda = 1$ is used. Meanwhile, exogenous variables are identified using a transfer function approach via cross-correlation function (CCF) analysis between exogenous and endogenous variables to determine the impulse response weights. The parameters (b, r, s) represent the lag time, decay pattern, and the number of coefficients before decay (Wei, 2006; Makridakis et al., 1998), which are identified based on the significance pattern of the CCF. In general, the GSTARX model can be expressed as (Pramoedyo et al., 2020):

$$\boldsymbol{Z}_t = \sum_{k=1}^{p} \sum_{l=0}^{\lambda_k} \boldsymbol{\Phi}_{kl} \boldsymbol{W}^{(l)} \boldsymbol{Z}_{t-k} + \omega(B)[\delta(B)]^{-1} \boldsymbol{D}_{t-b} + \boldsymbol{\varepsilon}_t.$$

Model parameters were estimated using the Seemingly Unrelated Regression (SUR) approach. Subsequently, diagnostic testing was carried out using the Ljung-Box test to assess the white noise assumption and the Mardia test to assess the multivariate normality assumption (Wei, 2006; Khufi et al., 2025).

The final stage of this study was to forecast using the obtained GSTARX model and to evaluate forecasting accuracy using Root Mean Square Error (RMSE) and Mean Absolute Percentage Error (MAPE), as described by Wei (2006).

$$RMSE = \sqrt{MSE} = \sqrt{\frac{1}{n} \sum_{t=1}^{n} \left(Z_i(t) - \hat{Z}_i(t) \right)^2},$$

$$MAPE = \frac{100\%}{n} \sum_{t=1}^{n} \left| \frac{Z_i(t) - \hat{Z}_i(t)}{Z_i(t)} \right|.$$

RESULTS AND DISCUSSION

Data Characteristics and Patterns Identification

Table 1. Descriptive Statistics

Region Police District	N	Mean	Std. Dev	Min	Max
Batu	60	20.45	8.28	5	40
Malang City	60	20.32	8.87	5	43
Malang	60	65.93	17.15	32	101
Pasuruan	60	65.58	21.58	27	126

Based on Table 1, this study uses 60 observations, derived from monthly data over a five-year period (2020 to 2024). In time series analysis, a dataset with more than 50 observations is generally considered sufficient to yield stable parameter estimates and effectively capture temporal patterns (Anwar & Rassiyanti, 2025). Therefore, the dataset used in this study is adequate to support the modeling process. Additionally, Batu Police District and Malang City Police District have a lower average claim frequency than Malang Police District and Pasuruan Police District. Pasuruan Police District also has the largest standard deviation, indicating higher fluctuations in claim frequency compared to other regions. The relatively wide range of minimum and maximum values across all regions indicates variation in claim frequency over time, particularly in Malang and Pasuruan Police Districts.

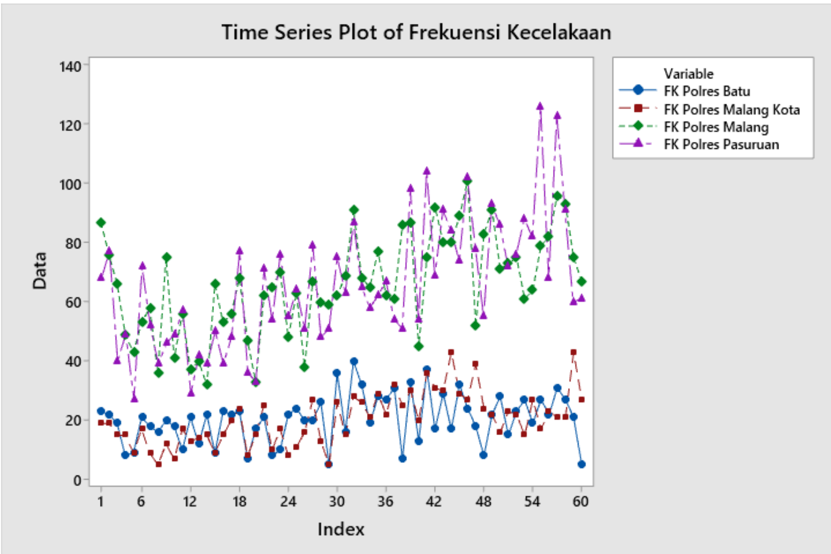

Figure 1. Time Series Plot

Based on Figure 1, the claim frequency across all police districts shows a fluctuating pattern throughout the observation period. Batu Police District and Malang City Police District have a relatively more stable pattern, while Malang Police District and Pasuruan

Police District show greater variation. Several simultaneous claim spikes occurring in more than one region indicate the presence of spatio-temporal relationships and possible influence from external factors.

Spatial Correlation

Table 2. Pearson Correlation

Region Police District	Batu	Malang City	Malang	Pasuruan
Batu p-value	1	0.390 (0.002)	0.347 (0.007)	0.540 (0.000)
Malang City p-value	0.390 (0.002)	1	0.520 (0.000)	0.498 (0.000)
Malang p-value	0.347 (0.007)	0.520 (0.000)	1	0.703 (0.000)
Pasuruan p-value	0.540 (0.000)	0.498 (0.000)	0.703 (0.000)	1

Based on Table 2, all regional pairs exhibit positive, statistically significant Pearson correlation coefficients at the 5% significance level. The highest correlation is observed between the Malang Police District and Pasuruan Police District, indicating a relatively strong spatial relationship between the two regions.

Data Stationarity in Variance and Mean

1. Stationarity in Variance

Table 3. Box-Cox Transformation

Region Police District	λ	Transformation (Z_t^*)		Conclusion
		Transformation Form	λ	
Batu	0.50	$\sqrt{Z_t}$	1.00	Stationary
Malang City	0.50	$\sqrt{Z_t}$	1.00	Stationary
Malang	1.00	-	-	Stationary
Pasuruan	0.00	$ln(Z_t)$	1.00	Stationary

Based on Table 3, the Box-Cox transformation results show that Batu Police District and Malang City Police District require a square root transformation. In contrast, Pasuruan Police District requires a natural logarithm transformation. At the same time, Malang Police District was already stationary in variance at the initial stage. After the transformation, all data met the stationarity assumption in variance.

2. Stationarity in Mean

Table 4. Augmented Dickey-Fuller (ADF) Test

Region Police District	p-value	Differencing p-value	Conclusion
Batu	0.000	-	Stationary
Malang City	0.509	0.043	Stationary
Malang	0.866	0.000	Stationary
Pasuruan	0.544	0.000	Stationary

Based on Table 4, the ADF test results show that only Batu Police District was stationary in mean at the initial stage, while the other regions required first-order differencing. After differencing, all data met the stationarity assumption in mean.

Cross Correlation Function (CCF)

The Cross Correlation Function is used to identify spatio-temporal relationships between regions by measuring the correlation between time series from different locations at various time lags. In this study, the CCF is specifically used to determine whether changes in claim frequency in one region are related to changes in other regions at the same time (lag 0), which serves as the basis for constructing location weights in the GSTARX model. The CCF significance boundary is determined using $\pm \frac{1,96}{\sqrt{T}}$ (Nurhayati et al., 2025). With $T = 59$, the significance boundary is ± 0.255, where a CCF value is considered significant if it exceeds this boundary.

Table 5. Inter-Regional CCF at Lag 0

Police District Pairs	CCF(0)	>0.255	Conclusion
Batu-Malang City	0.3949	Yes	Significant
Batu-Malang	0.0359	No	Not Significant
Batu-Pasuruan	0.2819	Yes	Significant
Malang City-Malang	0.2075	No	Not Significant
Malang City-Pasuruan	0.4049	Yes	Significant
Malang-Pasuruan	0.4957	Yes	Significant

Based on Table 5, most regional pairs have significant CCF(0) values, except for the Batu Police District-Malang Police District and Malang City Police District-Malang Police District pairs. Therefore, the formation of location weights only considers pairs with significant relationships.

GSTARX Model Identification

1. GSTAR Model Identification

Table 6. AIC Values for Candidate VAR Orders

VAR Order	AIC Value
VAR(1)	2.760126
VAR(2)	2.497717
VAR(3)	2.544284

Based on Table 6, the VAR(2) model has the smallest AIC value, so the selected time order is $p = 2$. With spatial order $\lambda = 1$, the GSTAR model used is GSTAR(2,1).

2. Exogenous Variable Identification

The CCF results between the calendar variation dummy variables and the claim frequency in each police district area did not show significant results, so the exogenous variable order used is $(b, r, s) = (0,0,0)$.

Thus, the GSTARX model used for all observation regions in this study is GSTARX(2,1)(0,0,0).

Thus, the GSTARX model used for all observation regions in this study is GSTARX(2,1)(0,0,0). This notation indicates that the model includes a temporal order of $p = 2$, representing two time lags, and a spatial order of $\lambda = 1$, representing first-order spatial dependence. The exogenous variable order $(b, r, s) = (0,0,0)$ indicates that no significant lagged effect of the exogenous variables is included in the model.

GSTARX Location Weights

This study employed two location weighting approaches: CCF(0)-based neighborhood weighting and cross-correlation normalized location weighting. Both of these weighting approaches are based on the cross-correlation function value at lag 0 (CCF(0)), and thus share a common foundation for representing spatial relationships between regions. The CCF(0)-based neighborhood weighting is determined binarily based on significant CCF values at lag 0: regions with significant relationships are assigned a weight of 1, and those without are assigned a weight of 0; thus, this approach directly identifies the presence of spatial relationships. Meanwhile, the cross-correlation normalization location weighting uses significant CCF(0) values as initial weights and assigns a weight of 0 to insignificant values, thereby not only considering the presence of relationships but also accounting for their strength across regions. The initial weight matrix is then normalized row-wise so that the sum of weights for each region equals 1. The normalized matrices for both location weighting approaches are shown below.

1. CCF(0)-Based Neighborhood Location Weights

$$W^{(1)} = \begin{bmatrix} 0 & 0.5 & 0 & 0.5 \\ 0.5 & 0 & 0 & 0.5 \\ 0 & 0 & 0 & 1 \\ 0.33 & 0.33 & 0.33 & 0 \end{bmatrix}$$

2. Cross-Correlation Normalization Location Weights

$$W^{(2)} = \begin{bmatrix} 0 & 0.583432 & 0 & 0.416568 \\ 0.493739 & 0 & 0 & 0.506261 \\ 0 & 0 & 0 & 1.000000 \\ 0.238421 & 0.342394 & 0.419186 & 0 \end{bmatrix}$$

By employing two approaches that share the same foundation but differ in their representations, this study aims to compare the effects of using binary weights and correlation-strength-based weights in capturing spatial interaction patterns, as well as their impact on the resulting modeling and forecasting outcomes. The selection of these two weights is based on the consideration that spatial relationships between regions need to be identified not only in terms of their existence but also in terms of the strength of their influence; thus, the use of these two approaches is expected to provide a more comprehensive representation of the structure of spatial interactions within the data.

GSTARX Model Parameter Estimation

Parameter estimation for the GSTARX(2,1)(0,0,0) model was performed using the SUR approach with two types of location weights.

1. CCF(0)-Based Neighborhood Location Weights
   a. Batu

      $Z_1(t) = 2.8048 + 0.3623Z_1(t-2).$

   b. Malang City

      $Z_2(t) = -0.7175Z_2(t-1) - 0.5130Z_2(t-2) + 0.5170\sum_{j\neq 2} W_{2j}^{(1)} Z_j(t-2).$

   c. Malang

      $Z_3(t) = -0.6846Z_3(t-1) - 0.4370Z_3(t-2).$

   d. Pasuruan

      $Z_4(t) = -0.8492Z_4(t-1) - 0.3323Z_4(t-2).$

2. Cross-Correlation Normalization Location Weights

a. Batu

$Z_1(t) = 2.7926 + 0.3605Z_1(t-2).$

b. Malang City

$Z_2(t) = -0.7216Z_2(t-1) - 0.5152Z_2(t-2) + 0.5219\sum_{j\neq 2} W_{2j}^{(2)} Z_j(t-2).$

c. Malang

$Z_3(t) = -0.6813Z_3(t-1) - 0.4318Z_3(t-2).$

d. Pasuruan

$Z_4(t) = -0.8527Z_4(t-1) - 0.3327Z_4(t-2).$

The parameter estimation results from both location-weight approaches yield relatively similar coefficient values. In Batu, Malang, and Pasuruan Police Districts, only the autoregressive component is significant. In contrast, in Malang City Police District, both the autoregressive and spatial components are significant. In general, the calendar variation dummy variables do not show a significant effect on claim frequency across the police district areas. This result may indicate that calendar effects do not primarily drive traffic accidents, but rather that factors such as traffic density, weather conditions, and driver behavior do. In addition, the variability of accident frequency over time may reduce the statistical significance of calendar effects. Furthermore, the spatial component is only significant in Malang City, which may reflect the characteristics of urban areas with higher mobility and stronger inter-regional interactions compared to other regions.

GSTARX Model Diagnostic Testing

1. White Noise Test

The Ljung-Box test was conducted on the model residuals for each police district area with test lag $m = 3$.

Table 7. Ljung-Box Test

Region Police District	Lag	CCF(0)-Based Neighborhood Location Weights		Cross-Correlation Normalization Location Weights	
		p-value	**Conclusion**	**p-value**	**Conclusion**
Batu	3	0.6867	White Noise	0.6898	White Noise
Malang City	3	0.3966	White Noise	0.3922	White Noise
Malang	3	0.2886	White Noise	0.2948	White Noise
Pasuruan	3	0.1516	White Noise	0.1429	White Noise

Based on Table 7, all police district areas under both location weight approaches have p-values > 0.05, indicating that the residuals contain no autocorrelation and are white noise.

2. Multivariate Normality Test

The Mardia test was conducted on the model residuals to test the multivariate normality assumption.

Table 8. Mardia Test

Test	CCF(0)-Based Neighborhood Location Weights		Cross-Correlation Normalization Location Weights	
	p-value	Conclusion	p-value	Conclusion
Mardia's Skewness	0.1820	Multivariate Normal	0.1765	Multivariate Normal
Mardia's Kurtosis	0.3598		0.3601	

Based on Table 8, the Mardia test results under both location weight approaches yield p-values > 0.05, indicating that the residuals follow a multivariate normal distribution. Accordingly, the GSTARX(2,1)(0,0,0) model with both location-weight approaches satisfies the model's diagnostic assumptions and is therefore suitable for forecasting.

GSTARX Model Forecasting

Forecasting was carried out using the GSTARX(2,1)(0,0,0) model with two location weight approaches. Because the model was built using data that had been transformed and first-order differenced, the forecast results at the model scale must be back-transformed to the original scale before comparison with actual data.

Table 9. GSTARX Model Forecasting Results

Region Police District	Month	Forecasting		Actual
		CCF(0)-Based Neighborhood Location Weights	Cross-Correlation Normalization Location Weights	
Batu	January	17.91	17.56	41
	February	15.09	15.33	18
	March	18.37	18.23	14
Malang City	January	26.35	26.34	25
	February	28.98	29.02	12
	March	27.84	27.74	32
Malang	January	82.57	82.45	136
	February	78.25	78.20	64
	March	78.15	78.16	74

Pasuruan	January	68.51	68.31	67
	February	64.08	64.07	72
	March	66.69	66.75	129

Based on Table 9, both location-weighting approaches yield relatively similar forecasts and capture the claim frequency pattern in some regions, although the model has not yet fully captured extreme claim frequency spikes in certain periods.

Table 10. Forecasting Accuracy Evaluation

Region Police District	**CCF(0)-Based Neighborhood Location Weights**		**Cross-Correlation Normalization Location Weights**	
	RMSE	**MAPE**	**RMSE**	**MAPE**
Batu	13.67	34.58	13.84	34.08
Malang City	10.12	53.28	10.16	53.50
Malang	32.02	22.39	32.08	22.40
Pasuruan	36.28	20.52	36.24	20.41
Average	**23.02**	**32.60**	**23.08**	**32.60**

Based on Table 10, both location-weighting approaches show relatively similar forecasting performance, with an average RMSE of approximately 23 and a MAPE of approximately 32%. However, these MAPE values indicate that the level of forecasting error remains quite high. The best accuracy levels were observed in the Pasuruan and Malang Police Districts, with MAPE values of around 20-22%, while the Malang City Police District had the highest MAPE of around 53%, indicating that the model has limitations in generating accurate predictions in that area. Practically speaking, these results show that the model is better at describing general patterns in claim frequency trends than at generating precise predictions. Moreover, the model's inability to capture extreme spikes in claim frequency may be attributed to its linear nature, which is less capable of accommodating sudden, irregular fluctuations caused by unexpected events.

CONCLUSION

The results of this study indicate that the GSTARX(2,1)(0,0,0) model can represent the frequency of traffic accident compensation claims across four police districts by capturing temporal and spatial relationships. The estimation results show that the autoregressive component has a significant influence in most regions. In contrast, the spatial component is significant only in the Malang City Police District, and the calendar variation dummy variables do not show a significant effect. The comparison of the two location weighting approaches yields relatively similar forecasting performance, as indicated by comparable RMSE and MAPE values. However, the average MAPE of approximately 32%, as well as MAPEs exceeding 50% in Malang City, indicate that forecasting accuracy remains limited, particularly in certain regions. Therefore, the model is better suited to describing general patterns of claim frequency than to producing highly accurate forecasts. Further model development is needed to improve forecasting performance, especially in capturing extreme fluctuations in claim frequency.

ACKNOWLEDGMENTS

The authors would like to express sincere gratitude to PT Jasa Raharja Malang Branch for its willingness to provide the data and information necessary for this study.

DECLARATIONS

Author Contribution	: AHRA:: Conceptualization, Data Collection, Data Processing, Formal Analysis, Methodology, Writing-Original Draft, and Editing; SR: Writing-Review & Editing, Validation, and Supervision.
Funding Statement	: This research was funded by personal funds.
Conflict of Interest	: The authors declare no conflict of interest.
Additional Information	: Additional information is available for this article.

REFERENCES

Anwar, R., & Rassiyanti, L. (2025). Analisis komparasi model peramalan prophet dan arima dalam memprediksi harga saham penutupan PT ANTM. *Lattice Journal: Journal of Mathematics Education and Applied*, *5*(1), 57–74. https://doi.org/10.30983/lattice.v5i1.9478

Aulia, L., Diana, A. F., & Ginanjar, A. (2025). Penerapan model regresi poisson untuk prediksi frekuensi klaim asuransi pada perusahaan asuransi jiwa. *Jurnal Riset Rumpun Matematika Dan Ilmu Pengetahuan Alam*, *4*(1), 160–172. https://doi.org/10.55606/jurrimipa.v4i1.5204

Badan Pusat Statistik. (2025). *Jumlah Kecelakaan, Korban Mati, Luka Berat, Luka Ringan, dan Kerugian Materi*. https://www.bps.go.id/id/statistics-table/2/NTEzIzI=/jumlah-kecelakaan--korban-mati--luka-berat--luka-ringan--dan-kerugian-materi.html

Fadlurohman, A., Utami, T. W., & Wasono, R. (2020). Generalized space time autoregressive modeling with variable exogenous (GSTAR-X) (Case Study: Inflation In Six Cities Of Central Java). *Prosiding Seminar Nasional Unimus*, *3*, 26–36. https://prosiding.unimus.ac.id/index.php/semnas/article/view/633

Go, B. A., & Mauliddin. (2024). Menghitung premi asuransi kecelakaan lalu lintas yang optimal berdasarkan besar klaim dan frekuensi klaim. *Lambda: Jurnal Ilmiah Pendidikan MIPA dan Aplikasinya*, *4*(1), 23–35. https://doi.org/10.58218/lambda.v4i1.843

Gustiasih, R., & Saputro, D. R. S. (2018). Model generalized space time autoregressive integrated dengan eror autoregressive conditional heteroscedastic (GSTARI-ARCH). *Prosiding Konferensi Nasional Penelitian Matematika Dan Pembelajarannya (KNPMP)*, 457–464. https://proceedings.ums.ac.id/knpmp/article/view/2050/2005

Hikmawati, F., Suhartono, & Prastyo, D. D. (2021). A Hybrid GSTARX-Jordan RNN model for forecasting space-time data with calendar variation effect. *Journal of Physics: Conference Series*, *1752*. https://doi.org/10.1088/1742-6596/1752/1/012013

Kementerian Keuangan Indonesia. (2017). *Peraturan Menteri Keuangan Nomor 15/PMK.010/2017 tentang Besar Santunan dan Iuran Wajib Dana Pertanggungan Wajib Kecelakaan Penumpang Alat Angkutan Umum di Darat, Sungai/Danau, Feri/Penyeberangan, Laut, dan Udara*. Berita Negara Tahun 2017 Nomor 276. Jakarta. https://peraturan.bpk.go.id/Details/112594/pmk-no-15pmk0102017

Kementerian Keuangan Indonesia. (2017). *Peraturan Menteri Keuangan Nomor 16/PMK.010/2017 tentang Besar Santunan dan Sumbangan Wajib Dana Kecelakaan Lalu Lintas Jalan*. Berita Negara Tahun 2017 Nomor 279. Jakarta. https://peraturan.bpk.go.id/Details/112596/pmk-no-16pmk0102017

Khufi, R. A., Gunawan, J. V., Nainggolan, E. Y., Nasrudin, M., & Trimono. (2025). Analisis jumlah tenaga kesehatan di jawa timur tahun 2022-2023 menggunakan metode RM-MANOVA satu arah. *Jurnal Riset Sistem Informasi*, *2*(3), 29–38. https://doi.org/https://doi.org/10.69714/2hhgzp35

Makridakis, S., Wheelwright, S. C., & Hyndman, R. J. (1998). *Forecasting: Methods and Applications* (3rd ed.). John Wiley & Sons.

Mubarak, R., Kamaroellah, A., & Suhartono. (2023). Forecasting money flow in east java using the generalized space-time autoregressive with exogenous variable method. *Iqtishadia: Jurnal Ekonomi Dan Perbankan Syariah*, *10*(2), 179–199. https://doi.org/10.19105/iqtishadia.v10i2.10462

Nurhayati, Hamidi, M. R., Mukhaiyar, U., & Sari, K. N. (2025). Cross-correlation analysis in evaluating spatio-temporal data dependence of climate variables through the GSTAR Model. *Jurnal Matematika, Statistika dan Komputasi*, *21*(3), 813–831. https://doi.org/10.20956/j.v21i3.43665

Pemerintah Pusat Indonesia. (1964). *Undang-Undang Republik Indonesia Nomor 33 Tahun 1964 tentang Dana Pertanggungan Wajib Kecelakaan Penumpang*. Lembaran Negara Tahun 1964 Nomor 137, Tambahan Lembaran Negara Nomor 2720. Jakarta. https://peraturan.bpk.go.id/Details/50393/uu-no-33-tahun-1964

Pemerintah Republik Indonesia. (1964). *Undang-Undang Republik Indonesia Nomor 34 Tahun 1964 tentang Dana Pertanggungan Wajib Kecelakaan Lalu Lintas Jalan*. Lembaran Negara Tahun 1964 Nomor 138, Tambahan Lembaran Negara Nomor 2721. Jakarta. https://peraturan.bpk.go.id/Details/50395/uu-no-34-tahun-1964

Pemerintah Pusat Indonesia. (2009). *Undang-Undang Republik Indonesia Nomor 22 Tahun 2009 tentang Lalu Lintas dan Angkutan Jalan*. Lembaran Negara Tahun 2009 Nomor 96, Tambahan Lembaran Negara Nomor 5025. Jakarta. https://peraturan.bpk.go.id/Details/38654

Pramoedyo, H., Ashari, A., & Fadliana, A. (2020). Interpolasi kriging dalam pemodelan gstar-sur dan gstarx-sur pada serangan hama penggerek buah kopi. *Media Statistika*, *13*(1), 25–35. https://doi.org/10.14710/medstat.13.1.25-35

PT Jasa Raharja. (2024). *Layanan Prima Kami*. https://www.jasaraharja.co.id/id/service

Rahmani, U. P., Nisa, K., & Hamsyiah, N. (2025). Penerapan model generalized space time autoregressive (gstar) pada data inflasi beberapa kota. *Sciencestatistics: Journal of Statistics, Probability, and Its Application*, *3*(1), 38–54. https://doi.org/10.24127/sciencestatistics.v3i1.7526

Sugiyono. (2007). *Statistika Untuk Penelitian*. Bandung: CV Alfabeta.

Suhartono, Gazali, M. M., & Prastyo, D. D. (2018). VARX and GSTARX models for forecasting currency inflow and outflow with multiple calendar variations effect. *Matematika*, 57–72. https://doi.org/10.11113/matematika.v34.n3.1139

Suryani, & Saputro, D. R. S. (2018). Estimasi parameter model generalizedspace time autoregressive (GSTAR) menggunakan metode generalized least square (GLS). *Prosiding Konferensi Nasional Penelitian Matematika Dan Pembelajarannya (KNPMP)*, 465–472. https://proceedings.ums.ac.id/knpmp/article/view/2051/2006

Wei, W. W. S. (2006). *Time Series Analysis: Univariate and Multivariate Methods* (2nd ed.). Addison Wesley.

Zhao, L., Li, Z., & Qu, L. (2022). Forecasting of Beijing PM2.5 with a Hybrid ARIMA model based on integrated aic and improved gs fixed-order methods and seasonal decomposition. *Heliyon*, *8*(12), 1–16. https://doi.org/10.1016/j.heliyon.2022.e12239