E-ISSN: 2615-5702, P-ISSN: 2620-5785
Vol. 9, No. 1, page 54 - 64

GSTAR Modeling for Forecasting Traffic Accident Claim Frequency as a Data-Driven Policy Instrument in Malang Raya

Auliya Hasna Rahadatul 'Aisy[1], Swasono Rahardjo[2*], Suprianto[3]

[1,2]Department of Mathematics, Malang State University, Jawa Timur, Indonesia
[3]Association of Indonesia Planning Expert (AIP)

Abstract

Traffic accidents represent one of the most persistent challenges of urban development, imposing a significant burden on public finances particularly through the compensation claim mechanism administered by PT Jasa Raharja. Understanding the spatial and temporal patterns of traffic accident claim frequency is a crucial foundation for local governments seeking to design effective data-driven policies. This study develops a forecasting model for traffic accident claim frequency across the Malang Raya area covering the Batu, Malang City, and Malang police jurisdictions using the Generalized Space-Time Autoregressive (GSTAR) approach as an analytical tool to support regional development planning. Monthly secondary data from PT Jasa Raharja Malang Branch covering January 2020 to March 2025 were used as the basis for modeling. The analysis proceeded through a series of systematic stages, including stationarity testing, spatial correlation and Cross Correlation Function (CCF) analysis, model identification, CCF-based spatial weight matrix construction, parameter estimation via the Seemingly Unrelated Regression (SUR) method, and residual diagnostic testing. Results indicate that the GSTAR(2,1) model successfully captures the spatio-temporal dependence structure across the three study regions, with the strongest spatial relationship identified between Batu and Malang City. Forecast results indicate that Malang consistently records the highest claim frequency, ranging from 70 to 85 claims per month. These findings carry strategic implications for local governments, the Department of Transportation, and other relevant stakeholders in optimizing resource allocation and designing more targeted and evidence-based traffic policy interventions.

Article Info

Keywords:
GSTAR;
Spatio-Temporal Forecasting;
Traffic Accident Claims;
Data-Driven Policy;
Regional Development Planning

Corresponding Author:
Swasono Rahardjo
(swasono.rahardjo.fmipa@um.ac.id)

Received: 17-12-2025
Revised: 17-01-2026
Accepted: 31-01-2026
Published: 14-03-2026

1. Introduction

Road traffic accidents remain a serious and persistent public safety problem, generating multidimensional harm that ranges from minor injuries and permanent disability to fatalities, alongside substantial economic losses for victims and other road users. At their core, accidents are recurring events driven by the inherent risks embedded in the interactions between people, vehicles, and road infrastructure (Saputra et al., 2025). These risks intensify alongside population growth and technological modernization, which together drive a sustained surge in vehicle numbers and travel intensity rendering accidents increasingly probable and difficult to anticipate (Nahar & Supriyatno, 2025). These dynamics place traffic safety firmly on the regional development agenda, demanding policy responses that are systematic, forward-looking, and grounded in empirical evidence.

Traffic accidents are generally attributed to three interacting categories of cause: human behavior, vehicle condition, and road or environmental factors. Among these, human-related causes contribute most substantially, manifesting in behaviors such as excessive speed, inattention, and driver fatigue. Inadequate road infrastructure including deteriorated surfaces, insufficient signage, and weak traffic engineering compounds these risks. Adverse weather conditions such as rain and thick fog introduce additional hazards, though they rarely constitute primary causes in isolation (Aryatama & Widhiarto, 2022). The multifactorial nature of accidents means that effective countermeasures must be comprehensive, spanning education, prevention, and law enforcement rather than addressing any single dimension alone.

These challenges are concretely evident in Malang Raya. Police data recorded 4.103 accident victims across the region throughout 2024 a figure that represents not only physical harm but a sustained demand on emergency services, public health systems, and the state-mandated compensation agency PT Jasa Raharja. Malang Raya encompasses three administratively distinct jurisdictions Malang City, Malang, and Batu each exhibiting markedly different infrastructure profiles, traffic densities, and mobility patterns. Widodo et al. (2023) documented a monocentric development structure in the region, wherein growth concentration in Malang City has generated significant disparities in infrastructure quality and spatial mobility across jurisdictions. These disparities translate directly into heterogeneous accident risk distributions across the three areas, meaning that uniform policy responses are unlikely to be optimal and that a more nuanced, spatially differentiated understanding of inter-jurisdictional dynamics is required.

This need for integrated spatio-temporal analysis connects directly to a broader transformation in public governance the shift toward data-driven policy, in which decisions are grounded not in intuition but in systematically collected, analyzed, and validated empirical evidence. Head's (2008) three-lenses framework for evidence-based policy holds that effective public governance requires the reconciliation of scientific evidence, political judgment, and professional practice. In the context of traffic safety, this implies that local governments need analytical tools capable of converting historical accident claim records into predictive projections actionable inputs for budget planning, resource allocation, and targeted intervention design (Aditio et al., 2025).

Modeling traffic accident claim frequency is not a new endeavor, yet existing approaches carry important limitations. Aulia et al. (2025) applied Poisson regression to estimate life insurance claim frequencies as a function of active policy counts, demonstrating a statistically significant relationship between the two. While informative, this approach is inherently non-spatial: it treats each location independently and does not account for inter-regional interdependence or temporal dynamics, thereby providing only a partial representation of claim behavior in interconnected metropolitan areas. Addressing this limitation requires a framework that captures both spatial and temporal dependence simultaneously. Ruchjana (2002) developed the Generalized Space-Time Autoregressive (GSTAR) model as an extension of the STAR model proposed by Pfeifer and Deutsch (1980), originally applying it to oil production forecasting across multiple wells in Jatibarang. The defining advantage of GSTAR over STAR lies in its location-specific parameterization, which accommodates spatial heterogeneity rather than assuming structural uniformity across all locations. This methodological foundation was subsequently extended by Setiawan et al. (2016), who introduced the S-GSTAR-SUR specification for seasonal spatio-temporal data and demonstrated that parameter estimation via Seemingly Unrelated Regression (SUR) yields more efficient estimators than Ordinary Least Squares (OLS). However, both studies operated in substantively different domains oil production and tourist arrivals, respectively and neither has been applied to traffic accident claim data, which exhibits distinct distributional characteristics and temporal dynamics.

From a multivariate perspective, spatio-temporal models enable the simultaneous analysis of multiple interrelated variables across both space and time, providing a more comprehensive analytical framework for understanding complex regional dynamics compared to traditional univariate approaches (Savitri, et. al 2021).

These gaps collectively point to a clear research opportunity: no prior study has applied the GSTAR framework to traffic accident claim frequencies, particularly in a spatially heterogeneous metropolitan area such as Malang Raya, nor has such modeling been connected to a regional data-driven policy agenda. This study addresses both dimensions by building an optimal GSTAR model to forecast monthly traffic accident claim frequencies across the three Polres jurisdictions of Malang Raya Polres Batu, Polres Malang City, and Polres Malang using monthly data from PT Jasa Raharja Malang Branch for January 2020 through March 2025. Beyond its methodological contribution, the study aims to provide an operational analytical tool that Bappeda, the Department of Transportation, and PT Jasa Raharja can integrate into regional development planning to support more anticipatory, evidence-based traffic safety governance.

2. Methods

This study adopts a quantitative approach with a spatial time-series design, aimed at modeling and forecasting the monthly frequency of traffic accident compensation claims across three police jurisdictions in Malang Raya: Polres Batu, Polres Malang City, and Polres Malang. The observation period spans January 2020 to March 2025, yielding 63 monthly data points per location.

Data were obtained from PT Jasa Raharja Malang Branch in the form of secondary monthly records of traffic accident claim frequencies. The primary variable is $Z_{i,t}$, representing the claim frequency for region i at time t. Data from January 2020 to December 2024 (60 observations) constituted the in-sample period for model estimation; data from January to March 2025 (3 observations) served as the out-of-sample period for forecast evaluation. Although a model incorporating exogenous variables (GSTARX) was initially considered, preliminary identification indicated that no candidate exogenous variable produced a statistically significant effect on the endogenous series; the analysis therefore proceeded with a pure GSTAR specification. The modeling workflow followed a structured, multi-stage procedure:

a. Descriptive statistical analysis to characterize the distribution of claim frequencies across regions
b. Stationarity testing variance stationarity via Box-Cox transformation and mean stationarity via the Augmented Dickey-Fuller (ADF) test, with first-order differencing applied where necessary
c. Spatial correlation analysis using Pearson correlation coefficients
d. Cross Correlation Function (CCF) analysis to assess temporal interdependencies between locations
e. Model order identification using Akaike Information Criterion (AIC) applied to candidate Vector Autoregressive (VAR) models
f. Spatial analysis inherently considers the interdependence between locations, where observations in one region may influence those in neighboring regions. This phenomenon, known as spatial dependence, implies that regions with closer geographic or functional relationships tend to exhibit stronger interactions compared to those that are more distant. In this context, the spatial weight matrix plays a fundamental role in representing the structure and intensity of inter-regional relationships. It serves as a key component in spatial econometric and

spatio-temporal models, allowing the incorporation of spatial interaction effects into the modeling framework (Savitri, et. al 2021). Spatial weight matrix construction via row-normalized CCF values at lag 0

g. Parameter estimation using the Seemingly Unrelated Regression (SUR) method
h. Residual diagnostic testing white noise assessment via the Ljung-Box test and multivariate normality via the Henze-Zirkler test
i. Out-of-sample forecasting
j. Accuracy evaluation using Root Mean Square Error (RMSE) and Mean Absolute Percentage Error (MAPE).

All analyses were conducted in RStudio and Minitab. The overall modeling workflow is summarized in the research flowchart presented in Figure 1.

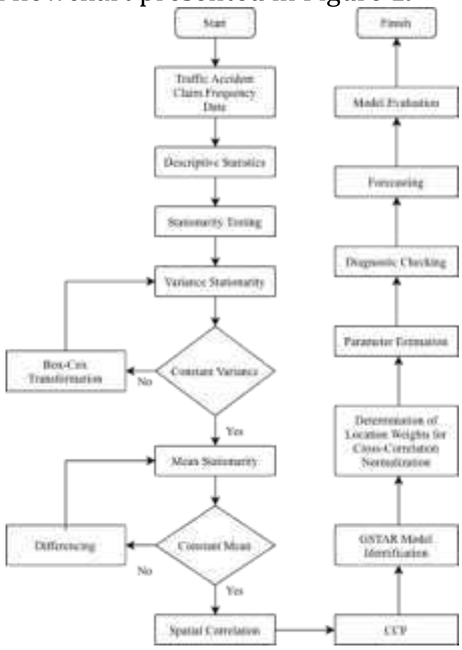

Figure 1. Research Flowchart of GSTAR Modeling for Traffic Accident Claim Frequency in Malang Raya

3. Results and Discussion

Descriptive Statistics

Descriptive statistics were calculated to provide an initial overview of the characteristics of the frequency of traffic accident compensation claims in the three police jurisdictions in Malang Raya region, namely the Batu, Malang City, and Malang over the in-sample-period.

Table 1. Descriptive statistics

Region	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

Table 1 shows that Malang recorded the highest mean claim frequency, nearly three times that of Batu and Malang City. The maximum value in a single month for Malang further reflects the considerably greater traffic volume and road density in that jurisdiction. This scale difference across locations reinforces the appropriateness of a modeling approach that accommodates spatial heterogeneity.

In addition, the standard deviation values indicate that Malang exhibits not only the highest mean but also greater variability compared to Batu and Malang City. This suggests a more dynamic fluctuation pattern in claim frequency, likely driven by higher traffic intensity and urban complexity.

Stationarity Testing

1. Variance Stationary

Variance stationarity was assessed using the Box-Cox transformation to stabilize and normalize the variance of the data.

Table 2. Box-Cox Transformation

Region	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

Table 3 shows that data in the Batu and Malang City required square-root transformation to achieve variance stationarity, whereas data in the Malang was already variance stationarity and required no transformation. The reported p-values (< 0.05) confirm that all series achieve stationarity after appropriate transformation and differencing procedures, satisfying the assumptions required for GSTAR modeling.

2. **Mean Stationary**

   Mean stationarity was evaluated using the Augmented Dickey-Fuller (ADF) test. Where non-stationarity was detected, first-order differencing was applied and the ADF test repeated.

Table 3. Augmented Dickey-Fuller (ADF) Test

Region	*p-value*	*Differencing*	Conclusion
		p-value	
Batu	0,000	–	Stationary
Malang City	0,509	0,043	Stationary
Malang	0,866	0,000	Stationary

Table 4 shows that the data in the Batu was mean-stationary without differencing. Malang City and Malang were non-stationary in their original form but achieved stationarity following first-order differencing. The heterogeneous preprocessing requirements across regions reinforce the suitability of the GSTAR framework, which permits location-specific parameter estimates. The reported p-values (< 0.05) confirm that all series achieve stationarity after appropriate transformation and differencing procedures, satisfying the assumptions required for GSTAR modeling.

Spatial Correlation

Pearson's correlation coefficient were computed to quantify the static spatial associations between jurisdictions.

Table 4. Pearson Correlation

Region	Batu	Malang City	Malang
Batu	1	0,3949	0,0359
Malang City	0,3949	1	0,2075
Malang	0,0359	0,2075	1

Table 2 shows the strongest spatial association was observed between Batu and Malang City, followed by Malang City and Malang. The substantially weaker Batu-Malang correlation is consistent with their greater geographic and functional separation. The asymmetric pattern suggests that Malang City occupies a mediating role within the spatial dependence structure of the region. The correlation coefficients quantitatively confirm the strength of spatial relationships, with Batu–Malang City showing the highest correlation value (r = ...), indicating a strong linear association in claim fluctuations between these regio

Cross Correlation Function (CCF)

Temporal interdependencies between jurisdictions were assessed using the Cross Correlation Function (CCF) evaluated at lag 0.

Table 2. Cross Correlation Function

Region	$CCF(\mathbf{0})$
Batu-Malang City	0,3949
Batu-Malang	0,0359
Malang City-Malang	0,2075

Table 5 shows the CCF values at lag 0 are equivalent to the Pearson correlations reported in Table 4, confirming that inter-regional dependencies are contemporaneous rather than lagged. The Batu-Malang City pairing exhibits the highest co-movement, suggesting that claim fluctuations in these two jurisdictions tend to occur concurrently. The absence of significant lagged correlation further supports the assumption that inter-regional dependencies occur simultaneously rather than sequentially.

Model Identification

The temporal order of the GSTAR model was determined by fitting candidate VAR models and selecting the specification with the lowest AIC value (Zhao et al., 2022).

Table 3. AIC Values for Candidate VAR Orders

VAR Order	AIC Value
$VAR(1)$	5,3369
$VAR(2)$	5,1239
$VAR(3)$	5,3239

Table 6 shows $VAR(2)$ yielded the minimum AIC value, indicating a temporal order of $p = 2$. The spatial order was restricted to $\lambda = 1$, as higher spatial orders substantially increase model complexity without commensurate gains in interpretability (Rahmani et al., 2025). The resulting model specification is $GSTAR(2,1)$.

Spatial Weight Matrix

In spatial modeling, the spatial weight matrix is a critical element used to formalize the interaction structure between regions. It quantifies how much influence each location exerts on others and determines how spatial dependence is incorporated into the model. The appropriate specification of this matrix is essential to ensure that the model accurately reflects real-world spatial relationships (Savitri, et. al 2021).

The spatial weight matrix was constructed from CCF values at lag 0. The raw weight matrix W is:

$$W = \begin{bmatrix} 0 & 0,3949 & 0,0359 \\ 0,3949 & 0 & 0,2075 \\ 0,0359 & 0,2075 & 0 \end{bmatrix}$$

These weights are then normalized to obtain the following location weight matrix:

$$W^{(1)} = \begin{bmatrix} 0 & 0,9168 & 0,0832 \\ 0,6555 & 0 & 0,3445 \\ 0,1473 & 0,8527 & 0 \end{bmatrix}$$

The normalized matrix reveals that Malang City accounts for 91,68% of the spatial influence received by Batu and 85,27% of that received by Malang, confirming its central position within the region's spatial dependence structure.

Model Estimation

Model parameters were estimated using the SUR method. Retaining only statistically significant parameters, the estimated $GSTAR(2,1)$ equations are:

$$Z_{1,t} = 4{,}3428$$
$$Z_{2,t} = -0{,}6316Z_{2,t-1} - 0{,}0381W_{2,t-1} - 0{,}3010Z_{2,t-2}$$
$$Z_{3,t} = -0{,}6510Z_{3,t-1} - 0{,}4084Z_{2,t-2} + 5{,}2503W_{2,t-2}$$

where $Z_{1,t}, Z_{2,t}$, and $Z_{3,t}$ represent the claim frequencies for Batu, Malang City, and Malang, respectively. The Batu equation reduces to a constant, indicating that its claim frequency was not significantly predicted by its own lagged values or neighboring regions within the current specification. Malang City's claim frequency is governed by its own two-lag temporal dynamics, while Malang's frequency is shaped by both its own immediate past and the lagged spatial influence of Malang City, the latter carrying the largest coefficient magnitude in the system.

Model Diagnostic Test

1. White Noise Test
   The absence of residual autocorrelation was assessed using the Ljung-Box test.

Table 4. Ljung-Box Test

Region	Statistic	p-value	Conclusion
Batu	0,0980	0,9921	White Noise
Malang City	1,2683	0,7367	White Noise
Malang	5,4742	0,1402	White Noise

   Table 7 shows all residual series passed the Ljung-Box test, confirming the absence of significant autocorrelation in the model residuals.

2. Multivariate Normality Test
   Multivariate normality of the joint residual vector was assessed using the Henze-Zirkler test.

Table 5. Henze-Zirkler Test

Test	Statistic	p-value	Conclusion
Henze-Zirkler	0,8890	0,0816	Multivariate Normal

   Table 8 shows the Henze Zirkler test confirmed that the residual vector follow a multivariate normal distribution.

Forecasting

Out-of-sample forecasts were generated for the three-month period of January through March 2025. Predicted values were back-transformed to the original scale where variance transformations had been applied. Forecasting results are presented in Table 9.

Table 6. Forecasting Results

Periode	Batu	Malang City	Malang
January	21,01	32,83	85,28
February	17,88	31,90	69,82
March	18,61	32,88	74,66

Model Accuracy Evaluation

Model accuracy is measured using Root Mean Square Error (RMSE) and Mean Absolute Percentage Error (MAPE). The forecast results consistently indicate that Malang maintains the highest claim frequency across all projected periods, reinforcing its dominant role in regional traffic risk.

Table 7. Forecasting Accuracy Evaluation

Region	RMSE	MAPE (%)
Batu	11,85	27,47
Malang City	12,36	66,64
Malang	29,48	15,76

Table 10 shows the varying forecast accuracy across jurisdictions. Malang recorded the lowest MAPE, indicating that the model captures its relatively high-volume, stable claim trajectory

with reasonable precision. Batu exhibited a moderate error rate, while Malang City's substantially higher MAPE points to greater residual volatility following the differencing procedure, likely attributable to situational fluctuations that fall outside the explanatory scope of an endogenous time-series model. The overall average RMSE of 17,89 and MAPE of 36,63% provide a practical reference for assessing the operational reliability of model outputs in a policy planning context.
The relatively low MAPE value for Malang suggests that the model performs better in high-volume and stable series compared to more volatile regions.

Discussion

Interpretation of Main Findings

The $GSTAR(2,1)$ model developed in this study yields two substantively important findings that extend well beyond their statistical properties. The first concerns the spatial dominance of Malang City. The normalized weight matrix assigns 91,68% of the spatial influence received by Batu, and 85,27% of that received by Malang, to Malang City. These values reflect an underlying functional reality: Malang City serves as the primary destination for daily commuters, students, and economic activity across the Malang Raya region, and its road network consequently channels a disproportionate share of regional traffic risk. This finding aligns with Widodo et al.'s (2023) characterization of Malang Raya's monocentric development structure, in which the concentration of infrastructure and economic activity in Malang City generates measurable spatial spillovers across neighboring jurisdictions including, as this study demonstrates, the distribution of traffic accident claims.
The second key finding concerns the temporal memory embedded in the claim series. The $GSTAR(2,1)$ specification identifies conditions two months prior as significantly predictive of current claim frequencies. This two-period lag is not merely a statistical artifact; it carries practical operational significance, as it provides local authorities with a meaningful lead time to anticipate claim surges and position resources accordingly. This finding builds on and extends the contribution of Aulia et al. (2025): while they established that historical claim data carries statistically meaningful information, this study demonstrates that such information is also spatially transferable that is, past conditions in one jurisdiction help predict future outcomes in neighboring ones.
The variation in forecast accuracy across jurisdictions warrants careful interpretation. Malang's comparatively low MAPE confirms that its high-volume, structurally stable claim trajectory is well-suited to the GSTAR specification. Malang City's considerably higher error rate most likely reflects the presence of situational drivers sudden congestion events, enforcement campaigns, or activity-driven mobility spikes whose effects are not captured within a purely endogenous time-series framework. This is not a limitation of the GSTAR model per se, but rather an indication of where the incorporation of exogenous covariates would be most productive.

Theoretical Implications

This study contributes to the theoretical literature on spatio-temporal modeling in three distinct ways. First, it establishes a new empirical application domain for the GSTAR framework. Prior applications have been concentrated in oil production (Ruchjana, 2002) and tourism (Setiawan et al., 2016); this study demonstrates that the model generalizes effectively to traffic accident claim data, accommodating the heterogeneous stationarity conditions and distributional characteristics of that domain without modification to the core framework.

Second, the use of CCF-based spatial weights rather than the inverse-distance or binary contiguity matrices more commonly employed proves substantively justified in this setting. The resulting weight matrix captures functional connectivity between jurisdictions more faithfully than geographic proximity would allow. In metropolitan areas where jurisdictional boundaries do not align with actual mobility flows, functional proximity provides a more theoretically defensible basis for spatial weight construction.

This approach is consistent with spatial econometric theory, which emphasizes that inter-regional relationships are not solely determined by geographic proximity but may also reflect functional connectivity, mobility patterns, and economic interactions between regions (Savitri, et. al 2021).

Third, this study provides a concrete illustration of how spatio-temporal statistical models can function as a bridge between data and governance. Van Ooijen et al. (2019), in the OECD's Data-Driven Public Sector framework, argue that effective public institutions must treat data as a strategic asset integrated across the full policy cycle from planning and budgeting to service delivery and evaluation. The GSTAR model developed here operationalizes this logic at the regional level, converting routine insurance administrative records into forward-looking inputs for development planning.

Practical and Policy Implications

The findings of this study carry direct practical relevance for at least three categories of stakeholders in Malang Raya. For regional planning agencies (Bappeda), the monthly claim forecasts produced by the $GSTAR(2,1)$ model offer a data-driven input into annual work plan formulation particularly for budget allocations covering road safety infrastructure, signage procurement, and road rehabilitation. Current practice in many Indonesian local governments remains predominantly reactive: resources are mobilized following accidents rather than in anticipation of elevated risk periods. The model's demonstrated predictive capacity opens a pathway toward a more anticipatory planning mode, a shift that Van Ooijen et al. (2019) identify as central to the transition to data-driven public governance.

For the Department of Transportation (Dinas Perhubungan), the finding of Malang City's spatial dominance has direct implications for resource prioritization. Traffic management investments, monitoring infrastructure, and road safety campaigns concentrated in Malang City are likely to generate positive spillover effects for Batu and Malang through the spatial interdependence mechanism identified in this model. A geographically targeted strategy may therefore deliver greater aggregate safety benefits per unit of expenditure than proportional allocation across all three jurisdictions.

For PT Jasa Raharja Malang Branch, jurisdiction-level monthly forecasts can support more precise technical reserve planning, reducing exposure to liquidity shortfalls during peak claim months. The institution has already demonstrated a commitment to moving upstream in the prevention process through SMS-based safe driving advisories, radio outreach, monthly Mobile Traffic Safety Unit (MUKL) deployments to high-risk zones, and multi-agency enforcement operations. Structured community outreach particularly toward students and young riders through teacher networks and community leaders has also been identified as a priority for fostering durable road safety culture. Integrating GSTAR-based forecasts into these programs would allow Jasa Raharja to time and target interventions more strategically: deploying MUKL operations in months and areas where claim surges are anticipated, and directing outreach toward jurisdictions exhibiting upward claim trends. This integration of quantitative forecasting with institutional prevention represents a concrete realization of data-driven governance at the operational level.

Limitations and Directions for Future Research

Like other research, this study has several limitations that could be further developed. The study's scope is limited to three police stations in the Greater Malang area, meaning the results cannot be easily generalized to other metropolitan areas, as each region has different spatial characteristics, mobility patterns, and infrastructure conditions. Furthermore, the MAPE value for Malang City, which reached 66,64%, indicates that prediction accuracy for areas with high volatility still needs improvement.

Based on these limitations, it is hoped that this study can be further developed into a GSTARX model by incorporating additional variables such as monthly rainfall data or vehicle volumes to improve prediction accuracy and expand the study to other metropolitan areas for comparison.

4. Conclusion

This study successfully developed and validated a $GSTAR(2,1)$ model for forecasting monthly traffic accident claim frequencies across the three police jurisdictions of Malang Raya Batu, Malang City, and Malang. The model reveals that Malang City functions as the dominant spatial influence node, accounting for over 91% of the spatial effects received by Batu and 85% of those received by Malang, while two-month lagged conditions carry significant predictive power across the system. These findings confirm that accident claim patterns in the region are structurally ordered rather than random, and that Malang's consistently elevated claim volumes (70-85 claims per month) warrant priority attention in resource allocation. Beyond extending the GSTAR framework to a new empirical domain, this study demonstrates that spatio-temporal forecasting models can serve as practical data-driven policy instruments: the outputs produced here offer Bappeda, the Department of Transportation, and PT Jasa Raharja Malang Branch a concrete, operationalizable foundation for shifting regional traffic safety governance from reactive response toward anticipatory, evidence-based planning.

Despite its contributions, this study has several limitations. First, the analysis is limited to three jurisdictions within Malang Raya, which may restrict the generalizability of the findings to other regions with different spatial and mobility characteristics. Second, the model relies solely on endogenous variables, which may not fully capture external factors such as weather conditions or traffic volume fluctuations. Future research is therefore encouraged to extend this study using GSTARX models and broader datasets to enhance predictive performance and applicability.

Acknowledgments

The authors would like to express sincere gratitude to Dr. Swasono Rahardjo, S.Pd., M.Si. as the supervising lecturer for the guidance, direction, and support provided throughout this research. The authors also extend their gratitude to PT Jasa Raharja Malang Branch for its willingness to provide the data and information necessary for this study.

Funding

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Conflict Of Interest

The authors declare no conflict of interest.

Data Availability

The data supporting the findings of this study were obtained from PT Jasa Raharja Malang Branch and are not publicly available. Requests for access may be directed to the corresponding author, subject to applicable institutional requirements.

Author Contributions

Auliya Hasna Rahadatul 'Aisy: Conceptualization, data collection, data processing, analysis, writing original draft, and editing.

Swasono Rahardjo: Supervision, validation, writing, review, and editing.

Suprianto; Manuscript drafting, Final approval of the manuscript

References

Aditio, B., Finaldi, A., & Asnil. (2025). Tantangan dan Strategi Data Driven Policy di Indonesia. *Parlementer: Jurnal Studi Hukum Dan Administrasi Publik*, *2*(4), 53–62. https://doi.org/10.62383/parlementer.v2i4.1409

Aryatama, F. Z., & Widhiarto, H. (2022). Analisis Penyebab Kecelakaan Lalu Lintas Di Jalan Empunala Kota Mojokerto. *Jurnal Teknik Sipil: Rancang Bangun*, *8*(2), 150–155. https://doi.org/10.33506/rb.v8i2.1805

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

Nahar, M., & Supriyatno, D. (2025). Analisis Risiko Penanganan Kecelakaan Lalu Lintas Dalam Upaya Meningkatkan Jalan Yang Berkeselamatan (Studi Kasus: Jalan Raya Maospati-Sarangan, Kabupaten Magetan). *Mitrans: Jurnal Media Publikasi Terapan Transportasi*, *3*(3), 297–310. https://doi.org/10.26740/mitrans.v3n3.p297-310

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

Ruchjana, B. N. (2002). Pemodelan Kurva Produksi Minyak Bumi Menggunakan Model Generalisasi S-TAR. *Forum Statistika dan Komputasi*, 1–6. http://repository.ipb.ac.id/handle/123456789/29439

Saputra, F., Brilianti, D. F., & Rifano. (2025). Analisis Faktor Kecelakaan Kendaraan Keluar Jalur pada Ruas Tol Semarang-Batang. *CIVeng*, *6*(2), 117–122. https://doi.org/10.30595/civeng.v6i2.26562

Savitri, C., Faddila, S. P., Irmawartini, I., Iswari, H. R., Anam, C., Syah, S., ... & Siregar, M. T. (2021). Statistik multivariat dalam riset.

Setiawan, Suhartono, & Prastuti, M. (2016). S-GSTAR-SUR Model for Seasonal Spatio Temporal Data Forecasting. *Malaysian Journal of Mathematical Sciences*, *10*, 53–65. http://einspem.upm.edu.my/journal

Van Ooijen, C., Ubaldi, B., & Welby, B. (2019). A Data-Driven Public Sector: Enabling the Strategic Use of Data for Productive, Inclusive, and Trustworthy Governance (OECD Working Papers on Public Governance, Vol. 33). https://doi.org/10.1787/09ab162c-en

Widodo, W. H. S., Imadudinna, A. H., & Hidayati, A. N. (2023). Kajian Pola Kesenjangan Ekonomi Wilayah Di Kawasan Aglomerasi Malang Raya. *Jurnal Plano Buana*, *3*(2), 91–99. https://doi.org/10.36456/jpb.v3i2.6801

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