BAREKENG: Journal of Mathematics and Its Applications
March 2026 Volume 20 Issue 1 Page 0069–0086
P-ISSN: 1978-7227 E-ISSN: 2615-3017
https://doi.org/10.30598/barekengvol20iss1pp0069-0086

# THE GENERALIZED SPACE-TIME ARIMA (GSTARIMA) MODEL FOR PREDICTING NITROGEN MONOXIDE TO MITIGATE EID AL- FITR AIR POLLUTION IN SURABAYA

**Hani Khaulasari [1*], Dian Candra Novitasari [2], Maunah Setyawati [3], Jeneiro Maulana [4], Shukor Sanim Mohd Fauzi [5]**

[1,2,3]*Mathematics Study Program, Faculty of Science and Technology, UIN Sunan Ampel Surabaya*
*Jln. Ahmad Yani No.117, Jemur Wonosari, Kec. Wonocolo, Surabaya, Jawa Timur, 60237, Indonesia*

[4]*Dinas Sumber Daya Air dan Bina Marga Kota Surabaya*
*Jln. Jimerto No.6, RW.8, Ketabang, Kec. Genteng, Surabaya, Jawa Timur, 60272, Indonesia*

[5]*College of Computing Informatics and Mathematics, University Technology MARA Perlis Branch*
*Perlis, 02600, Malaysia*

*Corresponding author's e-mail: * **hani.khaulasari@uinsa.ac.id***

***Article Info***

***Article History:***

*Received: 8th February 2025*
*Revised: 3rd May 2025*
*Accepted: 25th July 2025*
*Available online: 24th November 2025*

***Keywords:***

*Air Quality;*
*ARIMA;*
*GSTARIMA;*
*Inverted distance;*
*Mitigation;*
*Nitrogen monoxide;*
*Prediction;*
*RMSE;*
*sMAPE.*

***ABSTRACT***

*Air quality is a crucial factor due to its significant impact on environmental sustainability and public health. One of the major pollutants affecting air quality is Nitrogen Monoxide (NO), especially during periods of increased human mobility such as Eid al-Fitr. Monitoring and predicting NO levels are essential for early mitigation efforts. This study aims to evaluate the performance of the Generalized Space-Time Autoregressive Integrated Moving Average (GSTARIMA) model with three types of spatial weighting schemes and compare it with other forecasting methods, namely ARIMA, VARIMA, and Support Vector Regression (SVR), in predicting NO concentrations in Surabaya for April 2024. The data used in this study consist of daily NO concentration measurements obtained from the Surabaya City Environment Agency's monitoring stations located at SPKU Tandes, SPKU Wonorejo, and SPKU Kebonsari, covering the period from January 2023 to March 2024. The GSTARIMA model was selected for its capability to capture both spatial and temporal dependencies across monitoring locations. As an extension of the ARIMA model, GSTARIMA incorporates spatial weight matrices to model spatial heterogeneity. Parameter estimation was conducted using the Ordinary Least Squares (OLS) method. The results indicate that the GSTARIMA model with Inverse Distance Weighting (IDW) and order $(3,1,0)_1$ in the first spatial order yields the most accurate predictions, outperforming ARIMA, VARIMA, and SVR models. The model produced the lowest Symmetric Mean Absolute Percentage Error (sMAPE) of 0.93% and Root Mean Square Error (RMSE) of 5.32. A notable spike in NO concentrations was observed between April 23 and 25, 2024, coinciding with the post-Eid al-Fitr return flow, indicating a surge in population mobility.*



***How to cite this article:***

H. Khaulasari, D. C. Novitasari, M. Setyawati, J. Maulana and S. S. M. Fauzi., "THE GENERALIZED SPACE-TIME ARIMA (GSTARIMA) MODEL FOR PREDICTING NITROGEN MONOXIDE TO MITIGATE EID AL- FITR AIR POLLUTION IN SURABAYA," *BAREKENG: J. Math. & App.*, vol. 20, iss. 1, pp. 0069-0086, Mar, 2026.


Journal homepage: *https://ojs3.unpatti.ac.id/index.php/barekeng/*
Journal e-mail: *barekeng.math@yahoo.com; barekeng.journal@mail.unpatti.ac.id*

**Research Article · Open Access**

## 1. INTRODUCTION

Surabaya, with a population of around 3 million people, is predicted to continue to grow every year [1]. This increase in population is directly proportional to the growth in the number of motor vehicles, industrial activities, burning of fossil fuels, and an increase in garbage and waste, including open burning of garbage, which has an impact on the decline in air quality [2]. The amount of air pollution inhaled by the human body exceeds the standard and poses a proven health risk with a hazard quotient (HQ) of more than one and people living in urban industrial areas are at higher risk of respiratory disorders [3]. Air pollution gases such as PM2.5, SO2, and O3 have a direct impact on human health [4]. However, of the various types of pollutants present in the air, Nitrogen Monoxide (NO) gas is one of the main pollutants that affect the degradation of air quality in many areas of the earth's surface and is toxic [5]. Human activities that are not environmentally friendly have a positive correlation in producing prolonged air pollution emissions such as carbon dioxide (CO2), methane (CH4), and NO [6].

NO is a compound consisting of one nitrogen atom (N) and one oxygen atom (O) [7]. NO and nitrogen dioxide are known as Nitrogen Oxides (NOx) which are the major air pollutants and adversely affect the respiratory health of humans, animals, and the environment through the greenhouse effect and the depletion of the ozone layer [8]. NOx compounds are mostly produced by anthropogenic activities, which involve combustion processes such as energy production, transportation, and industrial activities [9].

NO air quality in Surabaya is influenced by various sectors, namely transportation (44%) as the largest contributor to air pollution, the energy industry sector (31%), the manufacturing industry (10%), the housing sector (14%), and the commercial sector (1%) also play a role in air pollution that damages health [10]. NO, as a major pollutant produced by motor vehicles and industry, can cause respiratory distress and increase the risk of heart disease [11]. Therefore, it is crucial to monitor NO levels in the air to reduce the adverse impact of pollution on public health [12]. Monitoring nitrogen levels in the air is essential to reduce the risk of adverse events. Air quality monitoring methods are based on time series analysis [13].

Time series analysis is a method used to understand, model, and predict data collected or observed in a specific time sequence [14]. Data in a time series consists of observations recorded sequentially at specific time intervals, such as daily, monthly, yearly, or even per second [15]. The main goal of the time series method is to identify patterns or trends in the data and utilize them to make predictions [16]. If the type of time series analysis involves a single variable measured in a specific time sequence using the univariate time series method approach [17] and univariate time methods such as ARIMA [18].

Research on univariate time series for air quality has been explored extensively. Abhilash et al. [19] used the ARIMA model to predict $NO_2$ levels, showing that stationary data can produce accurate predictions. Nieto [20], comparing four mathematical models namely Support Vector Machines (SVM), ARIMA, Multilayer Perceptron neural networks (MLPs), and Vector Autoregressive Moving-Average (VARMA) to predict PM10 pollutants in northern Spain, concluded that SVM outperformed others in short-term (one-month) and medium-term (seven-month) forecasts. Bernacki [21] combining the Random Forest algorithm with NASA's GEOS-CF data to predict $O_3$ and $NO_2$ concentrations in southeastern China for the next five days, has been shown to reduce prediction errors and prediction values in near real-time. In addition, Liu et al. [22] examined multivariate time series to predict NOx emissions using Vector Autoregressive (VAR) models, which captured the interactions among various pollutants and improved forecast accuracy.

Air quality monitoring stations are located at SPKU Wonorejo, SPKU Kebonsari, and SPKU Tandes. These locations reflect the diverse environmental characteristics of Surabaya and have the potential to influence air pollution levels [23]. SPKUA Wonorejo is located in an industrial and residential area with high economic activity. Kebonsari is a residential area affected by vehicle traffic, while SPKUA Tandes is situated in an area with industrial office land and business centers. These variations in environmental characteristics create diverse conditions, which can significantly affect air quality, as each location has its unique sources of pollution and environmental factors.

When time $(t)$ and location are combined and have a relationship between location and time, space-time series models such as the Generalized Space-Time Autoregressive Integrated Moving Average (GSTARIMA) model are used to analyze the data. This model is an extension of the Space-Time Autoregressive Integrated Moving Average (STARIMA) model [24]. The GSTARIMA model can capture

the relationship between a value at a given location and a value at a neighboring location at the same time, the relationship between a value at a given time and a random error at a previous time, and accommodate the difference in characteristics between different locations [25].

Various previous studies support the use of GSTARIMA in air quality monitoring. Research Hu et al. [26] predicted PM2.5 concentrations in 80 stations in the Beijing-Tianjin-Hebei region using the GSTAR model, which considers the effects of time and space. The prediction results of the GSTAR(1) model with inverted distance weights are superior to ARMA and STAR based on RMSE and MAE accuracy indicators. Akbar et al [25] concluded that the GSTARMAX($2_1$,$[7]_1$) model with inverted distance weights and OLS parameter estimation gave a small RMSE for SUF 1 (Prestasi Park), GSTAR(21) with SUR for SUF 6 (Wonorejo), and GSTARMA($2_1$,$[7]_1$) with SUR for SUF 7 (Kebonsari). The GSTARMA model can correct the prediction errors on CO data. Finally, Mohamed et al [27] using the GSTAR model to forecast the Air Pollution Index (API) in Selangor, considering the spatial-temporal relationship between locations. Models with inverted distance weights show the best performance with low RMSE, confirming the importance of location-based information in efficient forecasting. Other research related to GSTARIMA, namely Jamilatuzzahro et al [28] the GSTAR$(2,1)_I$ model with inverted distance weighting is the best model because it meets the assumptions of white noise and normal multivariate, and has the lowest RMSE and MAPE values. Imro'ah [29] concluded that the GSTAR(3; 1; 1) model with the MST weight matrix had a prediction error tolerance of 19%. Safira et al [30] resulting in the GSTARIMA model having the lowest average RMSE score compared to the ARIMA model. Ajobo et al. [24] shows that GSTARIMA with SUR estimation is more efficient in handling non-stationary data and correlates better with residual values.

This study identifies a gap in the literature that discusses air pollution predictions, especially in the Surabaya area with an emphasis on the Eid homecoming period. Most previous studies tended to focus on other major cities or use predictive models such as ARIMA and SARIMA, while the GSTARIMA model, which can map spatial and temporal relationships simultaneously, is still rarely applied. In addition, the impact of the Eid homecoming period on air pollution levels has not been widely investigated, although this has the potential to increase the number of vehicles and pollutant concentrations significantly. This study offers novelty by applying the GSTARIMA model to predict NO concentrations in three areas in Wonorejo, Kebonsari, and Tandes by considering specific spatial and temporal variations. The use of various types of weighting in the GSTARIMA model also enriches the analysis, resulting in more accurate and relevant predictions. This study aims to evaluate the best model of GSTARIMA with three spatial weightings (normalization of cross-correlation and uniform and inverse distance) and compare it with other methods such as ARIMA, VARIMA, and SVR and then predict changes in NO air pollution levels in Surabaya during Eid al-Fitr for the April 2024 period. The advantages of the method used can be further explained by its ability to capture both spatial and temporal dependencies more effectively than traditional models. The results of this modeling are expected to provide useful information for mitigating the impact of pollution and supporting the Surabaya Green and Clean Program, which aims to maintain the sustainability of the urban environment and improve the community's quality of life.

## 2. RESEARCH METHODS

The data used in this research are secondary data in the form of daily NO concentrations from January 2023 to March 2024, taken from monitoring by the Surabaya City Environment Agency (DLH) at three monitoring locations, namely at SPKU Wonorejo ($Y_1$), SPKU Kebonsari ($Y_2$), and SPKU Tandes ($Y_3$) Surabaya. The location of the monitoring is presented in Fig. 1. The air quality of NO concentration will be predicted using the GSTARIMA model using three weights ((i) normalized cross-correlation, (ii) uniform, and (iii) inverse distance), limited by the spatial order lag one. The steps involved in GSTARIMA modeling and forecasting are presented in Fig. 2. The type of instrument used, data collection approach, and data analysis techniques/tools can be detailed to ensure methodological transparency. Data processing and model estimation were carried out using Minitab and SAS software, which support comprehensive time series analysis and spatial modeling.

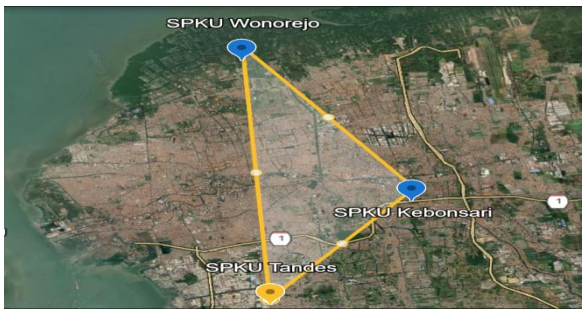



**Figure 1**. Air Monitoring Location in Surabaya

### 2.1 Generalized Space-Time Autoregressive Moving Average

STARIMA is a data modeling technique. The model establishes a relationship between time and location in observations. GSTARIMA, as an extension of the STARIMA model, was developed to overcome the limitations of the STARIMA model mainly related to the assumption of homogeneity of location characteristics [31]. Homogeneous location characteristics are difficult to find, as these traits often differ in research. The GSTARIMA model $(p, d, q)$ is a combination of time with the order of $p$ (autoregressive), the time component with the order $q$ (moving average), and the spatial component with the order of one. Essentially, the GSTARIMA model is very similar to the STARIMA model, only it allows for parameter values on a spatial lag that varies between locations [32]. The GSTARIMA model $(p, d, q)$ can be described in the form of [24], [25], [33] as follows Eq. (1).

$$\nabla \mathrm{Y_i}(\mathrm{t}^*) = \sum_{k=1}^{p} \sum_{l=0}^{\lambda} \Phi_{kl}\, \mathrm{W}^{(l)} \nabla \mathrm{Y_i}(\mathrm{t}^* - \mathrm{k}) - \sum_{k=1}^{q} \sum_{l=0}^{m} \Theta_{kl}\, \mathrm{W}^{(l)} \mathrm{a_i}(\mathrm{t}^* - \mathrm{k}) + a(\mathrm{t}^*), \tag{1}$$

where :

$\nabla \mathrm{Y_i}(\mathrm{t}^*)$ = $(1 - \mathrm{B})^d Y_i(\mathrm{t}^*)$ is an observation vector that has been stationary and is differencing at $t^* = 1, 2, \cdots, T$ at location $i = 1, 2, 3, \cdots, N$ size $(N \times 1)$;

$\Phi_{kl}$ = $diag\left(\emptyset_{kl}^1, \emptyset_{kl}^2, \cdots, \emptyset_{kl}^N\right)$ is the diagonal matrix of the autoregressive vector parameter at the $k$th lag and the $l$th spatial lag is sized $(N \times N)$;

$\Theta_{kl}$ = $diag\left(\theta_{kl}^1, \theta_{kl}^2, \ldots, \theta_{kl}^N\right)$ is the diagonal matrix of the vector moving average parameter at the $k$th lag and the $l$th spatial lag is sized $(N \times N)$.

$\mathrm{W}^{(l)}$ is a matrix of elements $W_{ij}^{(l)}$ spatial weights for spatial lag to $l$ sized $(N \times N)$ and when $(l = 0)$ then $\mathrm{W}^{(0)}$ is an identity matrix. $\mathrm{a_i}$ is a normally distributed error vector multivariate, $p$ is an autoregressive order, $q$ is the moving average order, $\lambda$ is the spatial order of autoregressive conditions and $m$ is the spatial order of the moving average condition. The estimation of the parameters of the GSTARIMA model using Ordinary Least Square is to minimize the number of squares of errors so that the estimate becomes $\hat{\beta} = (X^T \mathrm{X})^{-1} X^T y$.

The Modified Cross-Correlation Function (MCCF) and Modified Partial Cross-Correlation Function (MPCCF) matrix patterns can be used for data that have reached the stationary level. The plots of MCCF and MPCCF are depicted with the symbols (+), (−), and (·) in the $i$th and $j$th positions of the matrix. The MCCF plot determines the order of moving averages, which (cut off) after the order of zero at the lag of $q$, then the order of moving averages for the model is $MA(q)$. The value of $\hat{\rho}(\mathrm{k})$ indicated by the symbol (+) indicating a positive correlation between $i$ and $j$, the symbol (−) indicating a value less than -2 times the standard estimate of error, indicating a negative correlation between $i$ and $j$ and the symbol (·) indicating a value between -2 and 2 times the standard estimate of error, indicating that there was no correlation between $i$ and $j$ [34].

Residual check diagnostics consist of a white noise test and normal multivariate distribution. White noise multivariate testing is a technique to check whether a series of multivariate times has no autocorrelation in each variable and between variables. In other words, the data does not have a predictable temporal pattern. One of the commonly used tests for this purpose is the Ljung-Box multivariate test or the Portmanteau test [35], [36]. The hypothesis that $H_0$ is a residual White Noise and $H_1$ is a residual is not White Noise, with a

statistical test on Eq. (2), where $n$ is the number of observations and $R_k$ is the $k$th autocorrelation matrix. It is said to be residual white noise if the $p$-value is more than a significant level or $Q(h) < \chi^2_{n,\alpha}$ [37].

$$Q(h) = n^2 \sum_{k=1}^{h} \frac{1}{n} trace\left(R_k^T R_k\right). \tag{2}$$

In spatial-temporal modeling, spatial dependence is represented by a weighting matrix that reflects the influence among locations. The proper choice of weights is crucial to capture the true spatial structure of the data. This study uses three types of spatial weights, namely cross-correlation normalization (based on standardized correlations among locations), uniform weight (equal influence for all locations), and inverse distance weight (greater influence for closer locations).

1. Uniform Weights
   Uniform weighting is a technique in which each element in a dataset is given equal weight. This method calculates the average or aggregation of values across multiple locations assuming that each location has an equally large contribution. For example, if there are three locations, then the total neighbors $(n_i)$ is two [38].
2. Inverse Distance Weights (IDW)
   Inverse Distance weighting for location optimization is a technique that gives greater weight in closer locations and less weight in more distant locations. The weighting method applied in this spatial analysis aims to take into account the distance between geographical locations. The weight for interaction or relationship between two locations will be greater if the distance is closer. On the other hand, if the distance between locations is getting farther, the weight given will be smaller. Weight calculation is done using the inverse of the distance between two points or locations. This method is often used in spatial analysis to address the problem of spatial heterogeneity by maximizing the relative change of location based on distance. The calculation can use the distance in latitude ($u$) and longitude ($v$) coordinates between the centers of the observed location. Suppose we and represent the latitude and longitude coordinates of the location, and $d_{ij}$ represents the distance from the $i$-th location to the $j$-th location [33], [39].
   The Euclidean distance between locations is expressed as:
   $$d_{ij} = \sqrt{(u_i - u_j) + (v_i - v_j)}$$
   According to [40], the IDW can then be calculated using the following formula Eq. (3).
   $$w_{ij} = \frac{\frac{1}{d_{ij}}}{\sum_{j=1}^{n} \frac{1}{d_{ij}}}, \quad i \neq j. \tag{3}$$
3. Cross-Normalization Correlation Weights
   The method for normalizing weight values based on cross-variable or place cross-correlation is called cross-correlation normalization weighting. Each weight accurately reflects the relative contribution of the cross-correlation connections. This technique ensures that the total weight for each location or variable is equal to one. Cross-correlation between the $i$-th and $j$-th locations with the $k$-th time lag, where $(k)$ is the cross-correlation between the observations at the $i$-th and $j$-th locations at the $k$-th time lag. $\sigma_i$ and $\sigma_j$ are the standard deviations from the observations of the $i$-th and $j$-th locations. The weighting of the cross-correlation is as follows Eq. (4) [24].
   $$w_{ij} = \frac{r_{ij}(k)}{\sum_{i \neq j}^{n} |r_{ij}(k)|}, \quad i \neq j, \tag{4}$$
   with $r_{ij}(k) = \frac{\sum_{i,j=k+1}^{n} [Y_i(t^*) - \bar{Y}_i][Y_j(t^* - k) - \bar{Y}_j]}{\sqrt{\left(\sum_{i=1}^{n} (Y_i(t^*) - \bar{Y}_i)^2\right)\left(\sum_{i=1}^{n} (Y_j(t^*) - \bar{Y}_j)^2\right)}}$.

## 2.2 Forecasting Performance Model

Symmetric Mean Absolute Percentage Error (sMAPE) is an employed performance measure in this study. sMAPE describes a certain index for measuring an approximation model which is the average absolute

percentage error of the predicted value from the actual one [41]. As compared to the traditional mean absolute percentage error, sMAPE is more efficient at handling situations where the item for forecasting can assume the value of zero. In MAPE, some percentage of errors return undefined when the actual value is zero. sMAPE gives rise to errors that are valued between 0% to 200% unlike MAPE which can yield infinite values when actual results are nearing a figure close to indefinite [42]. This sMAPE eliminates those ambiguities making sMAPE more insightful and simpler to understand in different scenarios. sMAPE is also better for measuring the dynamic of data over time and thus understanding the state of the prediction model would be accurate and efficient even when the data is subjected to change [43]. One of the options for the performance criteria to resolve up against zero observation values is to apply sMAPE Eq. (5) [44] and Root Mean Square Error (RMSE) Eq. (6) [45].

$$sMAPE = \frac{1}{n}\sum_{t=1}^{n}\frac{\left|Y_{t^*} - \hat{Y}_{t^*}\right|}{\frac{\left(|Y_{t^*}| + |\hat{Y}_{t^*}|\right)}{2}} \times 100, \tag{5}$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{t=1}^{n}\left(Y_{t^*} - \hat{Y}_{t^*}\right)^2}. \tag{6}$$

where $Y_{t^*}$ is the actual value in the $t^*$-th period, $\hat{Y}_{t^*}$ is the prediction value in the $t^*$-th period, and $n$ is the number of periods used. Eq. (5) calculates the mean percentage of error by considering the predicted and actual values symmetrically, thus avoiding problems that can arise when the actual value is close to zero.

Order selection $(p, q)$ using Corrected Akaike Information Criterion (AICc) criteria with the formula Eq. (7) [46].

$$AICc = \left(2p - 2ln\left(L(\hat{\theta})\right)\right) + \frac{2p(p+1)}{n-p-1}, \tag{7}$$

where $p$ is the number of parameters in the model, and $L(\hat{\theta})$ is a function of likelihood.

### 2.3 Analysis Steps

The steps of the research analysis are as shown in Fig. 2.

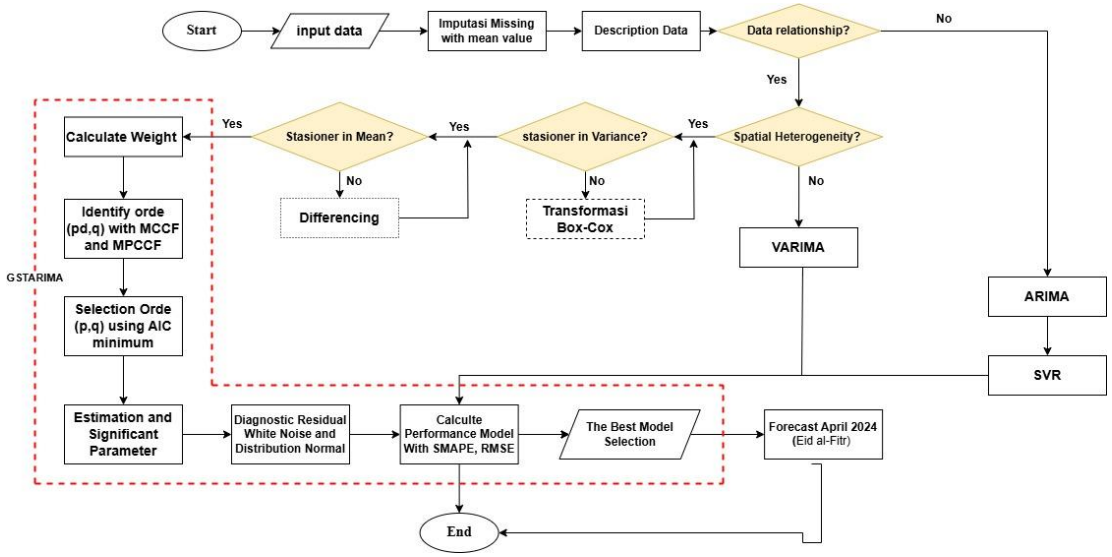



**Figure 2**. GSTARIMA Modeling and Forecasting Flowchart Scheme

1. Input Data on Nitrogen Levels at SPKU Wonorejo $(Y_1)$, SPKU Kebonsari $(Y_2)$, and SPKU Tandes $(Y_3)$.

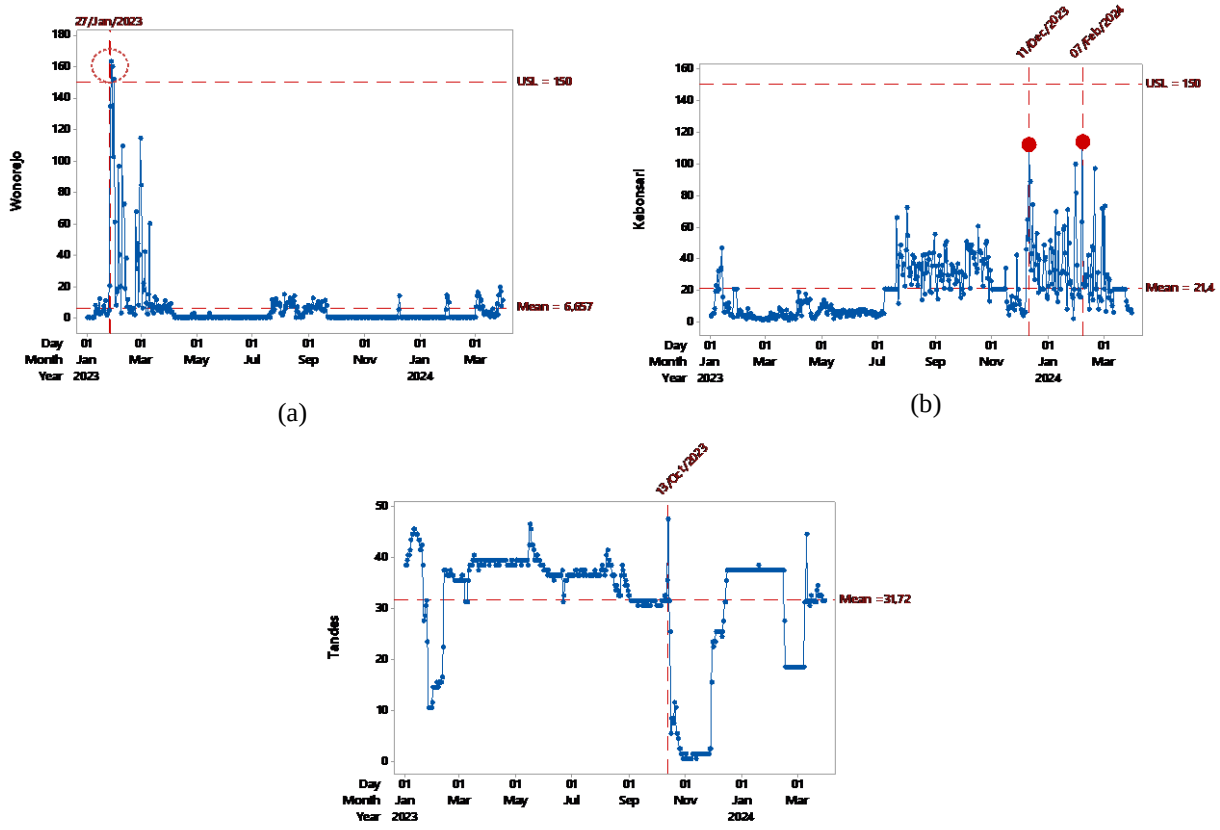
2. Description Time series plots were generated to examine trends, seasonal patterns, and extreme values. Descriptive statistics including the maximum and minimum values of NO concentrations were calculated and presented in Fig. 3.
3. Testing GSTARIMA Assumptions
   To validate the use of the GSTARIMA model, several preliminary tests were conducted:
   a. The Pearson correlation test using Equations of [47], to validate the use of the GSTARIMA model, several preliminary tests were conducted.
   b. Testing the Spatial Effects of Spatial Heterogeneity with the Breusch-Pagan Test using Equations of [48] and Spatial Dependency with the Morans'I Test using Equations of [49].
4. Detection of Stationary in variance and mean. If variance was non-stationary, a Box-Cox transformation was applied and If mean was non-stationary, differencing was used to stabilize the series. GSTARIMA modeling is a spatial-temporal analysis method that integrates the relationship between time and space. In order to construct a valid GSTARIMA model, the main stages need to be implemented systematically
   a. Calculating weights, this study uses weights (i) cross-correlation normalization weights using Eq. (4); (ii) Uniform weight using [38]; and (iii) weight inverse distance using Eq. (3)
   b. Identification of Lag Orders $(p, d, q)$ with Maximum Cross-Correlation Function (MCCF) and Modified Partial Cross-Correlation Function (MPCCF)**.**
   c. Selection order $(p, d, q)$ using Corrected Akaike Information Criterion (AICc) calculated with Eq. (7), and the model with the smallest AICc was selected.
   d. Model parameters were estimated using the Ordinary Least Squares (OLS).
   $$\hat{\beta} = \left(\mathrm{X}^{\mathrm{T}}\mathrm{X}\right)^{-1}\mathrm{X}^{\mathrm{T}}\mathrm{y}.$$
   Statistically, significant parameters were retained to construct the final GSTARIMA model equation.
   e. Diagnostic residual white noise using Eq. (2) and distribution normal
5. The predictive performance was assessed using sMAPE Eq. (5) and RMSE using Eq. (6). The model with the lowest values of sMAPE and RMSE was considered the best in predictive accuracy.
6. The best-performing GSTARIMA model was compared to three benchmark models are ARIMA, VARIMA, and Support Vector Regression (SVR). All models were trained using the same dataset, and their predictive performance was evaluated using the same metrics (sMAPE and RMSE).
7. Forecasting nitrogen levels in three locations in the April 2024 period to mitigate the impact of Eid al-Fitr in Surabaya based on the best GSTARIMA model.

## 3. RESULTS AND DISCUSSION

This study describes the distribution of variations in NO concentrations in three monitoring locations in Surabaya, namely SPKU Wonorejo, SPKU Kebonsari, and SPKU Tandes. Furthermore, an analysis was carried out using the GSTARIMA model to understand the spatial-temporal patterns of NO concentrations in the three regions. After the model is implemented, NO levels are forecasted for the April 2024 period to project changes in the concentration of this pollutant and provide deeper insights into air quality management in Surabaya.

### 3.1 Description of Nitrogen Monoxide

The air quality conditions of NO at SPKU Wonorejo, SPKU Kebonsari, and SPKU Tandes from January 1, 2023 to March 31, 2024 are depicted as shown in Fig. 3.




(a) (b) (c)

**Figure 3**. Description of Nitrogen Monoxide Concentration Data: (a) Wonorejo; (b) Kebonsari; (c) Tandes
*(Source: Data processing from Minitab)*

Based on Fig. 3, The nitrate concentration at the Wonorejo SPKU location generally ranges from 0 to 25.94 µg/m$^3$ with an average value of 6.6 µg/m$^3$, which falls under the "Good" air quality category. However, from January 26 to 30, 2023, there was a significant increase in NO concentrations, exceeding the acceptable threshold. The peak concentration was recorded on January 27, 2023, reaching 162.98 µg/m$^3$, categorized as "Unhealthy." This spike may be attributed to increased traffic congestion in the area, as Wonorejo is surrounded by university campuses and densely populated residential zones. During this period, a rise in vehicular movement was observed, possibly due to the return of students to Surabaya after the holiday break.

Meanwhile, Tandes, located in the western part of Surabaya, serves as a main access point to the Surabaya–Gresik toll road, facilitating both residential mobility and goods transportation. The Tandes area is dominated by residential and industrial areas. The concentration of NO in this area tends to spread in the interval 0 to 40.04 µg/m$^3$with an average value of 31.72 µg/m$^3$ which is included in the category of good air quality. On Friday, October 13, 2023, the highest NO concentration occurred, but overall, the NO concentration from January 1, 2023, to March 2024 has been within the specification limit.

Kebonsari is located in the south of the city of Surabaya. Kebonsari has a mix of urban and semi-urban environments and several green areas and parks that serve as the city's lungs. However, the area is also experiencing rapid urbanization. The concentration of NO spread from 0 to 43.24 µg/m$^3$ with an average of 31.72 µg/m$^3$. There were no observation points outside the threshold, so the concentration of NO was included in the category of good air quality. However, on Monday, December 11, 2023, Tuesday, January 30, 2024, and Wednesday, February 7, 2024, there was a spike in NO concentrations exceeding 100 µg/m$^3$ which is included in the unhealthy air quality category.

## 3.2 GSTARIMA Model

### 3.2.1 Correlation and Spatial Effect

The relationship between variables is measured using a correlation test, which indicates the presence or absence of a relationship and measures and quantifies the influence of one variable on another at a given time. This study employs Pearson's correlation analysis to examine NO air quality across three locations, as illustrated in Fig. 4. The correlation between the three locations is weak and negative, indicating an inverse relationship. This means that when NO levels in Wonorejo are high, they tend to be lower in Kebonsari and Tandes, and vice versa**.** Given this negative correlation, the GSTARIMA model is suitable for capturing the spatial and temporal dependencies among the locations.

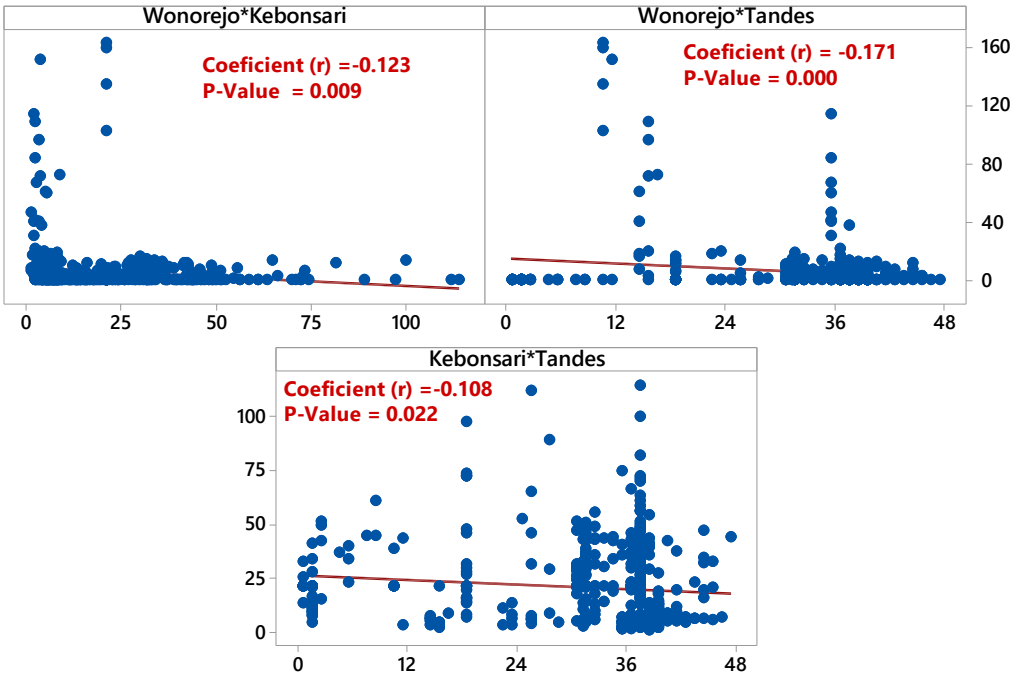



**Figure 4**. Scatterplot of Nitrogen Monoxide Concentrations at Three Locations
*(Source: Data processing from Minitab)*

Spatial heterogeneity was tested using the Breusch-Pagan test, while spatial dependency was assessed using Moran's I test. As a result, Table 1 showed that the significance value ($p$-value) of 0.045 and 0.032 was less than the significance level ($\alpha = 5\%$), then spatial heterogeneity occurred, namely the three locations had different characteristics that affect the air quality of NO.

**Table 1**. Spatial Effect Test

| Spatial Test | $p$-Value | Conclusion |
|---|---|---|
| Breusch- Pagan | 0.045 | Heterogeneity |
| Morans'I | 0.032 | Dependencies |

The main condition for using the GSTARIMA forecasting model is the existence of relationships in three locations, as shown in Fig. 4. If there is no relationship between nitrogen concentrations in the three locations, then the univariate time series model of ARIMA is used [50]. The second condition is that spatial heterogeneity occurs. If spatial homogeneity occurs, then use STARIMA [51]. If both conditions are met, the GSTARIMA model can be continued.

### 3.2.1 Detection of Stationary in Variance and Mean

The NO concentration at Wonorejo and Kebonsari was not stationary in variance ( $\lambda \neq 1$ ) as shown in column (a), unlike Tandes. After applying the Box-Cox transformation (column b), the variance became

stationary Table 2. After achieving variance stationarity, mean stationarity was tested using the MCCF plot as shown in Fig. 5 and the Dickey-Fuller (ADF) test shown in Table 3 can detect stationarity in the mean.

**Table 2**. Stationarity in Variance

| Nitrogen Monoxide of Location | Before Transformation (a) | After Transformation (b) |
|---|---|---|
| $Y_1$ (Wonorejo) | λ=1 | λ=1 |
| $Y_2$ (Kebonsari) | λ=1 | λ=1 |
| $Y_3$ (Tandes) | λ =0.5 | λ=1 |

Schematic Representation of Cross Correlations

| Variable/Lag | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Y1 | ++- | ++- | ++- | ++- | ++- | ++- | ++- | ++- | ++- | ++- | ++- |
| Y2 | ++- | ++- | ++- | ++- | ++- | ++- | ++- | ++- | ++- | ++- | ++- |
| Y3 | --+ | --+ | --+ | --+ | --+ | --+ | --+ | -.+ | -.+ | -.+ | ..+ |

\+ is > 2*std error, - is < -2*std error, . is between

(a)

Schematic Representation of Cross Correlations

| Variable/Lag | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Y1 | +-. | -.. | ... | ... | .+. | ... | ... | ... | ... | ... | ... |
| Y2 | -+. | .-. | .-. | ... | ... | ... | ... | .++ | .+. | .-. | ... |
| Y3 | ..+ | ..+ | ... | ..- | ... | ... | ... | ... | ... | -.. | +.. |

\+ is > 2*std error, - is < -2*std error, . is between

(b)

**Figure 5**. Detection of Mean Susceptibility with MCCF Plot: (a) The Data has not been Stationed in the Mean, (b) The Data has been Stationed in the Mean after Differencing
*(Source: Data processing from SAS)*

Fig. 5 shows an MCCF plot to detect stationarity in the mean, in plot (a) it can be seen that all locations in all lags contain different signs (+ and -) so that it is detected if the data is not stationary in the mean; (b) where at the time of the second lag has contained a sign (.) which means that the second lag in the 1st location and location 2nd is no longer out of the confidence interval or cut off in the second lag and the 5th lag of all locations there is no lag out of the confidence interval. The detection of stationarity in the mean can be confirmed through the Augmented Dickey-Fuller (ADF) test and the proposed GSTARIMA model with AR(3) and MA(1) with one-time differencing components is presented in Table 3.

**Table 3**. Dickey-Fuller Test of Nitrogen Monoxide in Three Locations after Differencing Once

| Location | $\tau_{count}$ | $p$-Value | $\tau_{456,5\%}$ | Results |
|---|---|---|---|---|
| $Y_1$ (Wonorejo) | -18.70 | <.0001 | -2.87 | Stationary in Mean |
| $Y_2$ (Kebonsari) | -22.17 | <.0001 | -2.87 | Stationary in Mean |
| $Y_3$ (Tandes) | -13.18 | <.0001 | -2.87 | Stationary in Mean |

In the test of the significance in the mean using the Dickey-Fuller test (ADF) after differencing once summarized in Table 3, the result was obtained that the ADF calculation value $(\tau_{hitung})$ in all three locations (Wonorejo, Kebonsari, and Tandes) is larger than the ADF Table $(\tau_{456,5\%})$ as -2.87 and the P-value value is less than the 5% significance level so it is concluded that the data has been stationary in the mean.

### 3.2.3 Spatial Weight Matrix

In this study, three types of spatial weights are used, namely (1) cross-correlation normalization weights, which are constructed based on the normalized cross-correlation values among locations, (2) uniform weight, where each location is assigned the same influence regardless of distance or correlation, and (3) inverse distance weight, where the influence between locations decreases as the distance between them increases.

1. Cross-correlation normalization

$$\mathbf{W}_{ij} = \begin{bmatrix} 0 & -0.0158 & 0.0188 \\ -0.0599 & 0 & 0.1189 \\ 0.1360 & -0.1824 & 0 \end{bmatrix}.$$

2. Uniform weights
It is divided into three locations so that there are two neighbors $n_i = 2$.

$$\mathbf{W}_{ij} = \begin{bmatrix} 0 & \frac{1}{2} & \frac{1}{2}0 \\ \frac{1}{2} & 0 & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0.50 & 0.50 \\ 0.50 & 0 & 0.50 \\ 0.50 & 0.50 & 0 \end{bmatrix}.$$

3. Inverse Distance Weight (IDW)
   IDW are calculated by taking into account the latitude and longitude coordinates of the locations. The latitude and longitude coordinates for each monitoring location are presented in Table 4 below.

**Table 4**. Latitude and Longitude Coordinates

| Location | Latitude (u) | Longitude (v) |
|---|---|---|
| Wonorejo | -7.2706 | 112.7128 |
| Kebonsari | -7.3271 | 112.6971 |
| Tandes | -7.2569 | 112.5994 |

Based on these coordinates, the spatial weighting matrix $\mathbf{W}_{ij}$ derived from the Inverse distance method is expressed as follows:

$$\mathbf{W}_{ij} = \begin{bmatrix} 0 & 0.6610 & 0.3390 \\ 0.6721 & 0 & 0.3279 \\ 0.5130 & 0.4870 & 0 \end{bmatrix}$$

Subsequently, the model identification was conducted using the Corrected Akaike Information Criterion (AICc). The minimum AICc values for different autoregressive (AR) and moving average (MA) orders are shown in Table 5.

**Table 5**. Minimum Information Criterion Based on AICC

| Lag | MA(q) 0 | MA(q) 1 |
|---|---|---|
| AR(p) 0 | -1.331638 | -1.487538 |
| AR(p) 1 | -1.432991 | -1.476406 |
| AR(p) 2 | -1.492226 | -1.502503 |
| AR(p) 3 | -1.524271* | -1.522911 |

*) Smallest AICC Value Selected Model

After selecting GSTARIMA $(3,1,0)_1$ as the best model based on the smallest AICc value, the parameters were estimated using the Ordinary Least Squares (OLS) method, and the results are shown in Table 6.

**Table 6**. Parameter Estimation of GSTARIMA $(3,1,0)_1$ Model

| Location | Parameter | Parameter Estimation | | |
|---|---|---|---|---|
| | | Cross-Normalized correlation (a) | Uniform (b) | Inverse Distance (c) |
| $Y_1$ (Wonorejo) | $\emptyset_{10}$ | 0.7013 | 0.6987 | 0.6991 |
| | $\emptyset_{11}$ | 0.0675 | 0.0614 | 0.0626 |
| | $\emptyset_{20}$ | -0.0240 | -0.0341 | -0.0335 |
| | $\emptyset_{21}$ | -2.7300 | -0.0482 | -0.0233 |
| | $\emptyset_{30}$ | 2.0700 | -0.0424 | -0.0422 |
| | $\emptyset_{31}$ | 1.9900 | -0.0387 | -0.0446 |
| $Y_2$ (Kebonsari) | $\emptyset_{10}$ | 0.5443 | 0.5406 | 0.5403 |
| | $\emptyset_{11}$ | -0.0448 | -0.0482 | -0.0480 |
| | $\emptyset_{20}$ | 0.2766 | 0.2674 | 0.2683 |
| | $\emptyset_{21}$ | -0.4130 | -0.0473 | -0.0302 |
| | $\emptyset_{30}$ | -0.2620 | 0.0250 | 0.0169 |
| | $\emptyset_{31}$ | 0.9210 | -0.1006 | -0.0911 |
| $Y_3$ (Tandes) | $\emptyset_{10}$ | 1.0706 | 1.0689 | 1.0686 |
| | $\emptyset_{11}$ | -0.0023 | 0.0012 | 0.0014 |
| | $\emptyset_{20}$ | -0.0977 | -0.0964 | -0.0961 |
| | $\emptyset_{21}$ | 0.0158 | 0.0010 | 0.0007 |

| Location | Parameter | Parameter Estimation | | |
|---|---|---|---|---|
| | | Cross-Normalized correlation (a) | Uniform (b) | Inverse Distance (c) |
| | $\emptyset_{30}$ | -0.0133 | 0.0109 | 0.0109 |
| | $\emptyset_{31}$ | 0.0400 | 0.0005 | 0.0011 |

Based on Table 5, the optimal selection of the GSTARIMA model is determined based on the lowest Corrected Akaike. The optimal model selected is GSTARIMA$(3,1,0)_1$ with $p$=3, $d$=1, $q$=0, and the first spatial lag, chosen due to its lowest AICc value. Parameter estimation using Cross-Normalized, Uniform, and Inverse Distance weights is presented in Table 6.

In addition, the GSTARIMA model is compared with other models for predicting NO concentrations to evaluate its effectiveness in capturing spatial and temporal dependencies. The comparison includes ARIMA(3,1,0), VARIMA(3,1,0), and SVR, assessed based on the average SMAPE and Root Mean Squared Error (RMSE), as shown in Table 7. This comparison evaluates GSTARIMA's predictive performance against alternative methods, emphasizing its advantages in modeling NO concentration across locations.

**Table 7**. Comparison of the Performance Model GSTARIMA, ARIMA, VARIMA and SVR

| Location | GSTARIMA $(3,1,0)_1$ Weight | | | ARIMA (3,1,0) | VARIMA (3,1,0) | SVR (RBF) C=10, ε=0.01, γ=0.1 |
|---|---|---|---|---|---|---|
| | Cross Normalized Correlation | Uniform | Inverse Distance | | | |
| $Y_1$ (Wonorejo) | 1.94% | 2.15% | 2.15% | 25.23% | 5.32% | 7.13% |
| $Y_2$ (Kebonsari) | 0.59% | 1.69% | 0.57% | 18.32% | 3.28% | 4.23% |
| $Y_3$ (Tandes) | 1.48% | 1.07% | 0.07% | 23.24% | 4.23% | 3.12% |
| Average sMAPE | 1.33% | 1.64% | **0.93%*** | 22.26% | 4.28% | 4.i3% |
| Average RMSE | 7.821 | 11.31 | **5.32*** | 25.23 | 18.37 | 16.23 |

*) The Best Performance Model Prediction

Based on the results shown in Table 7, The GSTARIMA(3,1,0) model excels in air quality forecasting with sMAPE below 10% and the lowest RMSE, demonstrating high accuracy. This indicates the existence of temporal and spatial dependency patterns in air pollution dynamics. This model effectively captures complex spatial effects in Tandes, Kebonsari, and Wonorejo, making it superior to models without spatial aspects such as ARIMA, VARIMA, or SVR models. In addition, an analysis of three spatial weighting methods showed that inverse distance weighting (IDW) gave the best results, with sMAPE of 0.93% and RMSE of 5.321. The advantage of this method lies in its ability to capture spatial influences between locations more effectively, where closer locations have greater influence weight. Thus, GSTARIMA with inverse distance weighting produces the most accurate predictions, making it the best choice for data analysis with complex spatial patterns.

The results of this study are in line with the research by [30]. Comparison of the GSTARIMA Model with ARIMA shows that the GSTARIMA model is the best. The best GSTARIMA model is an inverted distance weight matrix rather than uniform and cross-correlation spatial weighting, the goodness of the model is measured from an RMSE value of 0.184. The GSTARIMA model shows that the calculation of spatial dependence between provinces improves the ability of the inflation model to predict more accurately. In addition, in line with [38], modeling of provincial COVID-19 growth cases in Java on train passenger mobility is the best model with the IDW (inverted distance weight) matrix GSTAR $(1)_1$. Other than the study [27] showed that the GSTAR(1.1) model with inverted distance weights resulted in lower RMSE on average than the uniformly weighted model, with an increase in accuracy of 4.8%. Inverted distance weights are more effective in capturing spatial effects because they give higher priority to closer locations. Significant correlations between Air Pollutant Index (API) data in Banting, Petaling, and Shah Alam support the use of the GSTAR model for spatial-temporal analysis. Research by [52] also results in better-inverted distance spatial weighting than other spatial weightings with the GSTAR (1:1)-1(2) model, which is applied to consumer index price prediction. Then, the research by [53], shows that the GSTARIMA model (2,0,0) with an inverted distance weighting matrix achieves the lowest Mean Absolute Percentage Error (MAPE) in predicting flood events. Similarly, the studies by [33] and [54] demonstrate the effectiveness of GSTARIMA in forecasting rainfall and inflation, respectively.

However, the results of this study contradict previous studies [55], which showed that the best model was GSTAR (3,1,1) with MST weighting. On the contrary, the weighting of inverted distance gives a poor performance in predicting COVID-19 on the island of Java. The results of the white noise test Model GSTARIMA using IDW $(3,1,0)_1$ using the Portmanteau test obtained a P-value value of more than 5%. The results of the multivariate normal distribution test obtained a value of $d_{ij}^2$ more than $\chi^2_{(0.5)}$. The GSTARIMA modeling using Inverse Distance Weight types has met the residual White Noise and Normal Multivariate assumption so the GSTARIMA $(3,1,0)_1$ model is feasible.

### 3.3 Forecasting Nitrogen Monoxide Concentration Period April 2024

Forecasting the NO concentration in the April 2024 period using the inverse distance weighting selected as the best weighting can be seen in Fig. 6.

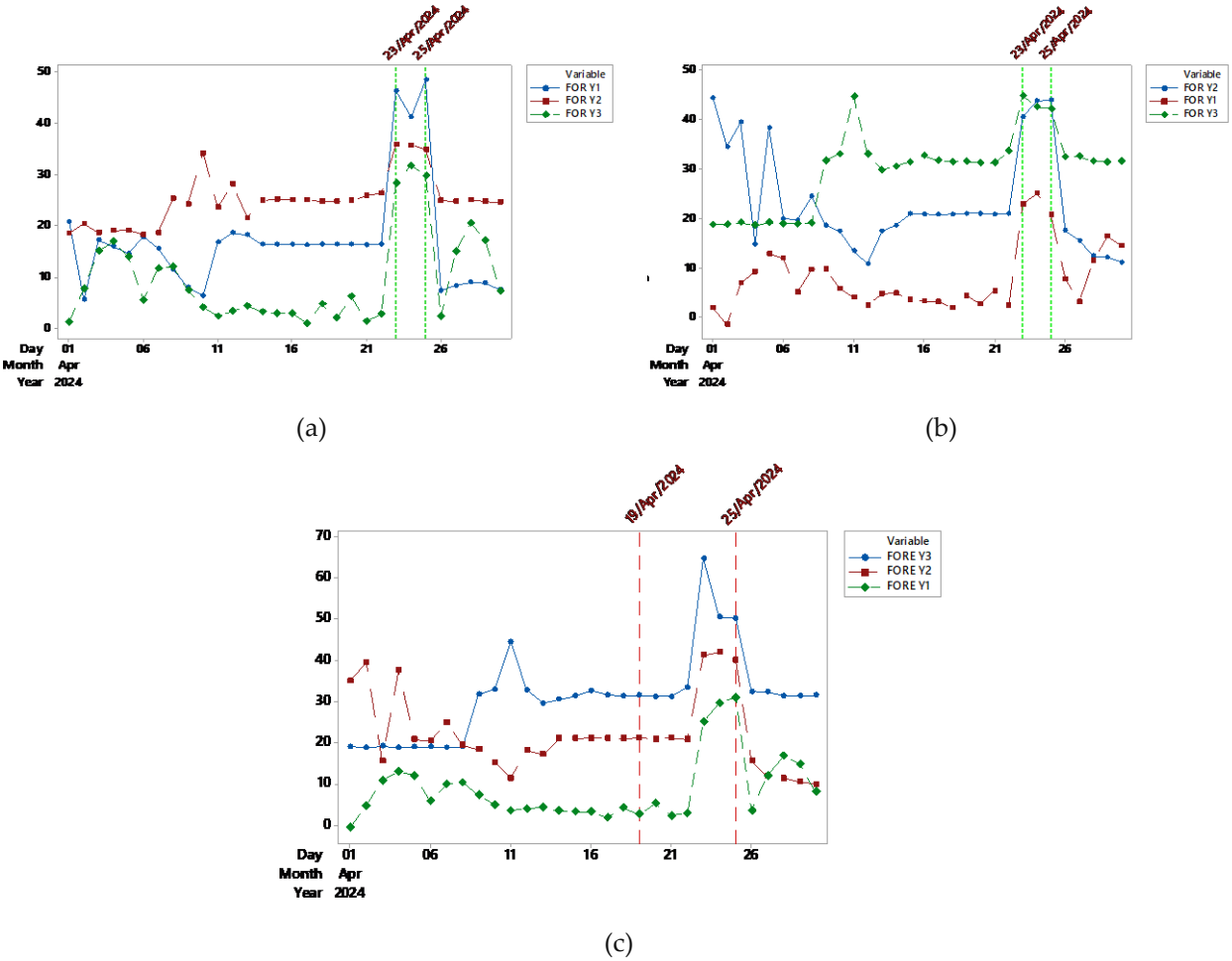



**Figure 6**. Forecasting the Quality of Nitrogen Monoxide Concentrations for the April 2024 Period on Three Types of Weighting: (a) Cross-Correlation Normalization Weights, (b) Uniform Weights, (c) Inverse Distance Weight
*(Source: Data processing from Minitab)*

In Fig. 6, it can be concluded that air quality forecasting based on NO concentrations in three monitoring locations, namely SPKU Wonorejo, Kebonsari, and Tandes, has been carried out using three types of spatial weighting. From the results of the analysis, it can be seen that the highest concentration of NO occurred from April 23 to 25, 2024. This period is included in the period of homecoming flows back to Surabaya after the Eid al-Fitr holiday which falls on April 21 to 22, 2024. During the Eid al-Fitr holiday, from April 19 to April 22, 2024, the concentration of NO in the air is relatively low. This is due to the reduction of activities in the city because many regional residents return to their hometowns to celebrate Eid al-Fitr with their families. This condition leads to a decrease in emissions from motor vehicles and industry, which is a major source of NO. NO levels are influenced by the NO levels of the previous day, namely one, two, three, and four days ago, at the three locations analyzed. In addition, there was a significant spatial

relationship between the three sites, suggesting that changes in NO levels in one location could affect NO levels in other locations.

However, when the return flow began to occur, from April 23 to 25, many residents returned to Surabaya. This increase in activity leads to a surge in the number of motor vehicles on the highway, which in turn increases NO emissions in the atmosphere. This phenomenon is reflected in air quality monitoring data that shows an increase in NO concentrations during the period. The increase in NO levels in Surabaya is also influenced by calendar variations, such as the Eid al-Fitr period, where the return flow of Eid al-Fitr is depicted as increasing nitrogen levels. In addition, New Year's celebrations also contribute to an increase in air pollution. Research [56] shows that pollutant concentrations have a significant increase ($p \leq 0.001$) compared to ordinary days. However, this study has not accommodated the effects of calendar variations. The addition of exogenous variables such as the addition of a dummy of calendar variations (for example holidays, weekends, or certain events) can be used as an alternative for further research, as has been done by [25], [57]. In addition, the use of GSTARIMA offers methodological strength because it is able to conduct simultaneous forecasting at various observation sites taking into account spatial and temporal dependencies, making it ideal for modeling the dynamics of pollutants that vary between space and time.

## 4. CONCLUSION

The best forecasting model used to predict Nitrogen Monoxide (NO) concentrations is GSTARIMA $(3,1,0)_1$ with a spatial order lag of one. Models with IDW help in capturing NO dispersion patterns more accurately. In the model evaluation, the GSTARIMA with IDW $(3,1,0)_1$ in the first spatial order provides a very accurate prediction of NO concentration compared to using other experimental prediction models (ARIMA, VARIMA, and SVR), with a very low sMAPE value of 0.93% and the lowest RMSE of 5.32.

The forecast results show that NO has the characteristic of fluctuating dispersion over time and between locations. However, there was a sharp spike in NO concentrations in the April 2024 period, especially from April 23 to 25, 2024. This surge was caused by the return flow of homecoming to Surabaya, which occurred after the holidays. This backflow increases the volume of vehicles entering the city, thereby increasing NO emissions from motor vehicles. This condition led to a significant increase in the concentration of NO in the air during that period.

Understanding NO concentration patterns and their influencing factors is essential for air quality management. This study establishes a foundation for spatial-temporal modeling using GSTARIMA, whose advantages, such as capturing spatial-temporal dependencies, flexible spatial weighting, and multi-location forecasting, can be compared with ARIMA, VARIMA, and SVR. Future research may include exogenous variables (e.g., calendar dummies, temperature, humidity) in extended models like GSTARIMA-X [58], [32]. Additionally, outlier detection should be considered, as previous studies [59]. showed improved performance when integrated into GSTAR. Finally, if residual correlations exist across locations, the GSTAR-SUR model [60] could be employed using Seemingly Unrelated Regression for more accurate estimation.

## Author Contributions

Hani Khaulasari: Conceptualization,Formal Analysis, Methodology ,Visualization,Writing - Review and Editing. Dian Candra Novitasari: Data curation, Resources, Validation. Maunah Setyawati: Writing - Original Draft, Project Administration. Jeneiro Maulana: Project Administration, Software. Shukor Sanim Mohd Fauzi: Supervision, Resources. All authors discussed the results and contributed to the final manuscript.

## Funding Statement

This study is an outcome of a research grant funded by the Institute for Research and Community Service (LPPM) of UIN Sunan Ampel Surabaya. The author expresses sincere gratitude for the support provided.

## Acknowledgment

The researcher gratefully acknowledges LPPM UIN Sunan Ampel Surabaya for funding support, the Surabaya Environmental Agency (DLH) for data provision, and the reviewers for their valuable suggestions and corrections.

## Declarations

This research is carried out jointly according to the division of tasks of each one, without any conflict between authors

## REFERENCES

[1] BPS Kota Surabaya, *KOTA SURABAYA DALAM ANGKA 2023*. Surabaya: BPS Kota Surabaya, 2023. [Online]. Available: https://surabayakota.bps.go.id/

[2] R. Kurniawan et al., "IMPACTS OF INDUSTRIAL PRODUCTION AND AIR QUALITY BY REMOTE SENSING ON NITROGEN DIOXIDE CONCENTRATION AND RELATED EFFECTS: AN ECONOMETRIC APPROACH", *Environmental Pollution*, vol. 334, Oct. 2023, doi: https://doi.org/10.1016/j.envpol.2023.122212.

[3] T. Nakyai, M. Santasnachok, A. Thetkathuek, and N. Phatrabuddha, "INFLUENCE OF METEOROLOGICAL FACTORS ON AIR POLLUTION AND HEALTH RISKS: A COMPARATIVE ANALYSIS OF INDUSTRIAL AND URBAN AREAS IN CHONBURI PROVINCE, THAILAND", *Environmental Advances*, vol. 19, Apr. 2025, doi: https://doi.org/10.1016/j.envadv.2024.100608.

[4] R. Zhao, X. Huang, J. Xue, and X. Guan, "A PRACTICAL SIMULATION OF CARBON SINK CALCULATION FOR URBAN BUILDINGS: A CASE STUDY OF ZHENGZHOU IN CHINA", *Sustainable Cities and Society*, vol. 99, Dec. 2023, doi: https://doi.org/10.1016/j.scs.2023.104980.

[5] S. Mukherjee, G. Kalra, and S. C. Bhatla, "ATMOSPHERIC NITROGEN OXIDES (NOX), HYDROGEN SULPHIDE (H2S) AND CARBON MONOXIDE (CO): BOON OR BANE FOR PLANT METABOLISM AND DEVELOPMENT?", *Environmental Pollution*, Jan. 2025, doi: https://doi.org/10.1016/j.envpol.2025.125676.

[6] D. Z. L. Benjamin and B. Dieudonné, "A REVIEW OF THE EFFECTS OF CLIMATE CHANGE ON HYDROPOWER DAMS IN CAMEROON", *Journal of Environmental & Earth Sciences*, vol. 6, no. 3, Sep. 2024, doi: https://doi.org/10.30564/jees.v6i3.6735.

[7] M. Dai, B. Gu, X. Ma, and T. Chun, "NITROGEN MONOXIDE REDUCTION BY CARBON MONOXIDE TO COMBUSTION CONTROL WITH CALCIUM FERRITE REDOX IN IRON ORE SINTERING", *Fuel*, vol. 337, Apr. 2023, doi: https://doi.org/10.1016/j.fuel.2022.127172.

[8] S. Srivastava, S. Behera, and D. Mitra, "DISTRIBUTION OF OZONE, CARBON MONOXIDE AND OXIDES OF NITROGEN OVER AN URBAN LOCATION IN THE FOOTHILLS OF THE NORTH-WESTERN HIMALAYAS", *Urban Climate*, vol. 55, May 2024, doi: https://doi.org/10.1016/j.uclim.2024.101913.

[9] D. Cubides, X. Guimerà, I. Jubany, and X. Gamisans, "A REVIEW: BIOLOGICAL TECHNOLOGIES FOR NITROGEN MONOXIDE ABATEMENt", *Chemosphere*, vol. 311, Jan. 2023, doi: https://doi.org/10.1016/j.chemosphere.2022.137147.

[10] A. D. Sakti et al., "MULTI-AIR POLLUTION RISK ASSESSMENT IN SOUTHEAST ASIA REGION USING INTEGRATED REMOTE SENSING AND SOCIO-ECONOMIC DATA PRODUCTS", *Science of The Total Environment*, vol. 854, Jan. 2023, doi: https://doi.org/10.1016/j.scitotenv.2022.158825.

[11] S. A. Meo, M. A. Salih, J. M. Alkhalifah, A. H. Alsomali, and A. A. Almushawah, "ENVIRONMENTAL POLLUTANTS PARTICULATE MATTER (PM2.5, PM10), CARBON MONOXIDE (CO), NITROGEN DIOXIDE (NO2), SULFUR DIOXIDE (SO2), AND OZONE (O3) IMPACT ON LUNG FUNCTIONS", *Journal of King Saud University - Science*, vol. 36, No. 7, Aug. 2024, doi: https://doi.org/10.1016/j.jksus.2024.103280.

[12] T. Salthammer, "CARBON MONOXIDE AS AN INDICATOR OF INDOOR AIR QUALITY", *Environmental Science: Atmospheres*, vol. 4, No. 3, pp. 291–305, Mar. 2024, doi: https://doi.org/10.1039/D4EA00006D.

[13] M. N. A. Ramadan, M. A. H. Ali, S. Y. Khoo, M. Alkhedher, and M. Alherbawi, "REAL-TIME IOT-POWERED AI SYSTEM FOR MONITORING AND FORECASTING OF AIR POLLUTION IN INDUSTRIAL ENVIRONMENT", *Ecotoxicology and Environmental Safety*, vol. 283, Sep. 2024, doi: https://doi.org/10.1016/j.ecoenv.2024.116856.

[14] P. S. Thorat *Et Al.*, "ON THE TIME SERIES ANALYSIS OF RESISTIVE SWITCHING DEVICES", *Microelectronic Engineering*, vol. 297, Mar. 2025, doi: https://doi.org/10.1016/j.mee.2024.112306.

[15] Z. Shu and pp. W. Chan, "APPLICATION OF FRACTAL ANALYSIS ON WIND SPEED TIME SERIES: A REVIEW", *Advances in Wind Engineering*, Jan. 2025, doi: https://doi.org/10.1016/j.awe.2024.100028.

[16] F. A. Chyon, Md. N. H. Suman, Md. R. I. Fahim, and Md. S. Ahmmed, "TIME SERIES ANALYSIS AND PREDICTING COVID-19 AFFECTED PATIENTS BY ARIMA MODEL USING MACHINE LEARNING", *Journal Of Virological Methods*, vol. 301, Mar. 2022, doi: https://doi.org/10.1016/j.jviromet.2021.114433.

[17] K. G. Ranjan, B. R. Prusty, and D. Jena, "REVIEW OF PREPROCESSING METHODS FOR UNIVARIATE VOLATILE TIME-SERIES IN POWER SYSTEM APPLICATIONS", *Electric Power Systems Research*, vol. 191, Feb. 2021, doi: https://doi.org/10.1016/j.epsr.2020.106885.

[18] A. Al-Lami and Á. Török, "REGIONAL FORECASTING OF DRIVING FORCES OF CO2 EMISSIONS OF TRANSPORTATION IN CENTRAL EUROPE: AN ARIMA-BASED APPROACH", *Energy Reports*, vol. 13, pp. 1215–1224, Jun. 2025, doi: https://doi.org/10.1016/j.egyr.2025.01.004.

[19] M. S. K. Abhilash, A. Thakur, D. Gupta, and B. Sreevidya, "TIME SERIES ANALYSIS OF AIR POLLUTION IN BENGALURU USING ARIMA MODEL", *Ambient Communications and Computer Systems*, 2018, pp. 413–426. doi: https://doi.org/10.1007/978-981-10-7386-1_36.
[20] P. J. García Nieto, F. Sánchez Lasheras, E. García-Gonzalo, and F. J. De Cos Juez, "PM10 CONCENTRATION FORECASTING IN THE METROPOLITAN AREA OF OVIEDO (NORTHERN SPAIN) USING MODELS BASED ON SVM, MLP, VARMA AND ARIMA: A CASE STUDY", *Science Of The Total Environment*, vol. 621, pp. 753–761, Apr. 2018, doi: https://doi.org/10.1016/j.scitotenv.2017.11.291.
[21] J. Bernacki, "FORECASTING THE AIR POLLUTION CONCENTRATION WITH NEURAL NETWORKS", *Urban Climate*, vol. 59, pp. 102262, Feb. 2025, doi: https://doi.org/10.1016/j.uclim.2024.102262.
[22] J. Liu *Et Al.*, "MACHINE LEARNING FOR FORECASTING FACTORY CONCENTRATIONS OF NITROGEN OXIDES FROM UNIVARIATE DATA EXPLOITING TREND ATTRIBUTES", *International Journal Of Advanced Nuclear Reactor Design and Technology*, vol. 6, no. 2, pp. 117–122, Jun. 2024, doi: https://doi.org/10.1016/j.jandt.2024.12.002.
[23] Dinas Lingkungan Hidup Surabaya, "PEMKOT SURABAYA PANTAU KUALITAS UDARA DENGAN INDEKS PENCEMARAN UDARA MENGGUNAKAN 5 PARAMETER". [Online]. Available: Https://Www.Surabaya.Go.Id/
[24] S. Ajobo, O. O. Alaba, and A. Zaenal, "GENERALISED SPACE-TIME SEASONAL AUTOREGRESSIVE INTEGRATED MOVING AVERAGE SEEMINGLY UNRELATED REGRESSION MODELLING OF SEASONAL AND NON-STATIONARY DATA", *Scientific African*, vol. 24, pp. E02189, Jun. 2024, doi: https://doi.org/10.1016/j.sciaf.2024.e02189.
[25] M. Akbar, B. Ruchjana, D. Prastyo, A. Muhaimin, and E. Setyowati, "A GENERALIZED SPACE-TIME AUTOREGRESSIVE MOVING AVERAGE (GSTARMA) MODEL FOR FORECASTING AIR POLLUTANT IN SURABAYA", Presented At The Journal Of Physics: Conference Series, IOP Publishing, 2020, pp. 012022, doi: https://doi.org/10.1088/1742-6596/1490/1/012022
[26] J. Hu, S. Wang, and J. Mao, "SHORT TIME PM2.5 PREDICTION MODEL FOR BEIJING-TIANJIN-HEBEI REGION BASED ON GENERALIZED SPACE TIME AUTOREGRESSIVE (GSTAR)", *IOP Conf. Ser.: Earth Environ. Sci.*, vol. 358, no. 2, Dec. 2019, doi: https://doi.org/10.1088/1755-1315/358/2/022075.
[27] N. M. Mohamed, N. H. A. Rahman, and H. S. Zulkafli, "GENERALIZED SPACE-TIME AUTOREGRESSIVE (GSTAR) FOR FORECASTING AIR POLLUTANT INDEX IN SELANGOR", *Journal of Quality Measurement and Analysis*, vol. 19, no. 3, pp. 143–153, 2023.
[28] R. E. Jamilatuzzahro, R. Caraka, A. S. Herliansyah, D. M. Sari, and B. Pardamean, "GENERALIZED SPACE TIME AUTOREGRESSIVE OF CHILI PRICES", in *2018 International Conference on Information Management and Technology (Icimtech)*, Sep. 2018, pp. 291–296. doi: https://doi.org/10.1109/ICIMTech.2018.8528117.
[29] N. Imro'ah, "DETERMINATION OF THE BEST WEIGHT MATRIX FOR THE GENERALIZED SPACE TIME AUTOREGRESSIVE (GSTAR) MODEL IN THE COVID-19 CASE ON JAVA ISLAND, INDONESIA", *Spatial Statistics*, vol. 54, 2023, doi: https://doi.org/10.1016/j.spasta.2023.100734
[30] A. Safira, R. A. Dhiya'ulhaq, I. Fahmiyah, and M. Ghani, "SPATIAL IMPACT ON INFLATION OF JAVA ISLAND PREDICTION USING AUTOREGRESSIVE INTEGRATED MOVING AVERAGE (ARIMA) AND GENERALIZED SPACE-TIME ARIMA (GSTARIMA)", *Methodsx*, vol. 13, pp. 102867, Dec. 2024, doi: https://doi.org/10.1016/j.mex.2024.102867.
[31] W. W. S. Wei, *TIME SERIES ANALYSIS UNIVARIATE AND MULTIVARIATE METHODS*, Second Edition. Boston San Francisco , New York: Pearson, Addison Wesley, 2006.
[32] A. Ashari, A. Efendi, and H. Pramoedyo, "GSTARX-SUR MODELING USING INVERSE DISTANCE WEIGHTED MATRIX AND QUEEN CONTIGUITY WEIGHTED MATRIX FOR FORECASTING COCOA BLACK POD ATTACK IN TRENGGALEK REGENCY", Presented At The IISS 2019: Proceedings Of The 13th International Interdisciplinary Studies Seminar, IISS 2019, 30-31 October 2019, Malang, Indonesia, European Alliance For Innovation, 2020, pp. 50.doi: https://doi.org/10.4108/eai.23-10-2019.2293086
[33] N. Ilmi, A. Aswi, and M. K. Aidid, "GENERALIZED SPACE-TIME AUTOREGRESSIVE INTEGRATED MOVING AVERAGE (GSTARIMA) DALAM PERAMALAN DATA CURAH HUJAN DI KOTA MAKASSAR", *Inferensi*, vol. 6, no. 1, pp. 25–43, 2023.doi: https://doi.org/10.12962/j27213862.v6i1.14347
[34] W. W. S. Wei, *MULTIVARIATE TIME SERIES ANALYSIS AND APPLICATIONS*. in Wiley Series in Probability and Statistics. New Jersey: John Wiley & Sons Ltd, 2019.
[35] M. Li and Y. Zhang, "BOOTSTRAPPING MULTIVARIATE PORTMANTEAU TESTS FOR VECTOR AUTOREGRESSIVE MODELS WITH WEAK ASSUMPTIONS ON ERRORS", *Computational Statistics & Data Analysis*, vol. 165, Jan. 2022, doi: https://doi.org/10.1016/j.csda.2021.107321.
[36] M. Hallin and H. Liu, "CENTER-OUTWARD RANK- AND SIGN-BASED VARMA PORTMANTEAU TESTS: CHITTURI, HOSKING, AND LI–MCLEOD REVISITED", *Econometrics and Statistics*, Feb. 2023, doi: https://doi.org/10.1016/j.ecosta.2023.01.006.
[37] Š. Hudecová and M. Šiman, "STOCHASTIC HYPERPLANE-BASED RANKS AND THEIR USE IN MULTIVARIATE PORTMANTEAU TESTS", *Journal Of Multivariate Analysis*, Jun. 2024, doi: https://doi.org/10.1016/j.jmva.2024.105344.
[38] U. S. Pasaribu, U. Mukhaiyar, N. M. Huda, K. N. Sari, and S. W. Indratno, "MODELLING COVID-19 GROWTH CASES OF PROVINCES IN JAVA ISLAND BY MODIFIED SPATIAL WEIGHT MATRIX GSTAR THROUGH RAILROAD PASSENGER"S MOBILITY", *Heliyon*, vol. 7, no. 2, Feb. 2021, doi: https://doi.org/10.1016/j.heliyon.2021.e06025.
[39] Sifriyani, I. N. Budiantara, M. F. F. Mardianto, and Asnita, "DETERMINATION OF THE BEST GEOGRAPHIC WEIGHTED FUNCTION AND ESTIMATION OF SPATIO TEMPORAL MODEL – GEOGRAPHICALLY WEIGHTED PANEL REGRESSION USING WEIGHTED LEAST SQUARE", *Methodsx*, vol. 12, Jun. 2024, doi: https://doi.org/10.1016/j.mex.2024.102605.
[40] A. J. Hama Rash, L. Khodakarami, D. A. Muhedin, M. I. Hamakareem, and H. F. H. Ali, "SPATIAL MODELING OF GEOTECHNICAL SOIL PARAMETERS: INTEGRATING GROUND-BASED DATA, RS TECHNIQUE, SPATIAL

STATISTICS AND GWR MODEL", *Journal Of Engineering Research*, vol. 12, no. 1, pp. 75–85, Mar. 2024, doi: https://doi.org/10.1016/j.jer.2023.10.026.
[41] W. Yang, S. N. Sparrow, and D. C. H. Wallom, "A COMPARATIVE CLIMATE-RESILIENT ENERGY DESIGN: WILDFIRE RESILIENT LOAD FORECASTING MODEL USING MULTI-FACTOR DEEP LEARNING METHODS", *Applied Energy*, vol. 368, Aug. 2024, doi: https://doi.org/10.1016/j.apenergy.2024.123365.
[42] L. C. Velasco, A. J. Estose, M. Opon, E. Tabanao, and F. Apdian, "PERFORMANCE EVALUATION OF SUPPORT VECTOR REGRESSION MACHINE MODELS IN WATER LEVEL FORECASTING", *Procedia Computer Science*, vol. 234, pp. 436–447, Jan. 2024, doi: https://doi.org/10.1016/j.procs.2024.03.025.
[43] C. García-Aroca, M. Asunción Martínez-Mayoral, J. Morales-Socuéllamos, and J. V. Segura-Heras, "AN ALGORITHM FOR AUTOMATIC SELECTION AND COMBINATION OF FORECAST MODELS", *Expert Systems With Applications*, vol. 237, Mar. 2024, doi: https://doi.org/10.1016/j.eswa.2023.121636.
[44] M. Effendi, D. D. Prastyo, and M. S. Akbar, "MODELING AND FORECASTING RETURN VOLATILITIES OF INTER-CAPITAL MARKET INDICES USING GARCH-FRACTIONAL COINTEGRATION MODEL VARIATION", *Procedia Computer Science*, vol. 234, pp. 389–396, Jan. 2024, doi: https://doi.org/10.1016/j.procs.2024.03.0199.
[45] C. Wang, F. Xie, J. Yan, and Y. Xia, "A U-MIDAS MODELING FRAMEWORK FOR FORECASTING CARBON DIOXIDE EMISSIONS BASED ON LSTM NETWORK AND LASSO REGRESSION", *Energy Reports*, vol. 13, pp. 16–26, Jun. 2025, doi: https://doi.org/10.1016/j.egyr.2024.11.069.
[46] F. Yanuar et al., "BAYESIAN ESTIMATION UNDER DIFFERENT LOSS FUNCTIONS FOR THE CASE OF INVERSE RAYLEIGH DISTRIBUTION", *Kuwait Journal Of Science*, vol. 52, no. 1, Jan. 2025, doi: https://doi.org/10.1016/j.kjs.2024.100343.
[47] H. Rahadian, S. Bandong, A. Widyotriatmo, and E. Joelianto, "IMAGE ENCODING SELECTION BASED ON PEARSON CORRELATION COEFFICIENT FOR TIME SERIES ANOMALY DETECTION", *Alexandria Engineering Journal*, vol. 82, pp. 304–322, Nov. 2023, doi: https://doi.org/10.1016/j.aej.2023.09.070.
[48] S. M. Yuni, T. M. Saputra, and N. N. Fadhilah, "THE IMPLEMENTATION OF GEOGRAPHICALLY WEIGHTED REGRESSION (GWR) METHOD ON OPEN UNEMPLOYMENT RATE IN REGENCY/CITY OF SUMATRA ISLAND", *BAREKENG: Jurnal Ilmu Matematika Dan Terapan*, vol. 19, no. 1, Jan. 2025, doi: https://doi.org/10.30598/barekengvol19iss1pp73-86.
[49] Y. Farida, M. Farmita, pp. K. Intan, H. Khaulasari, and A. T. Wibowo, "MODELING CRIME IN EAST JAVA USING SPATIAL DURBIN MODEL REGRESSION", *BAREKENG: Jurnal Ilmu Matematika Dan Terapan*, vol. 18, no. 3, Jul. 2024, doi: https://doi.org/10.30598/barekengvol18iss3pp1497-15088.
[50] F. Tolesh and S. Biloshchytska, "FORECASTING INTERNATIONAL MIGRATION IN KAZAKHSTAN USING ARIMA MODELS", *Procedia Computer Science*, vol. 231, pp. 176–183, Jan. 2024, doi: https://doi.org/10.1016/j.procs.2023.12.190.
[51] O. M. De Barros, C. L. Marte, C. A. Isler, L. R. Yoshioka, and E. S. Da Fonseca Junior, "SPATIAL MATRICES FOR SHORT-TERM TRAFFIC FORECASTING BASED ON TIME SERIES", *Latin American Transport Studies*, vol. 1, Dec. 2023, doi: https://doi.org/10.1016/j.latran.2023.100007.
[52] A. Zaki, L. Shafruddin, and I. Thaha, "APPLICATION OF THE GENERALIZED SPACE TIME AUTOREGRESSIVE (GSTAR) METHOD IN FORECASTING THE CONSUMER PRICE INDEX IN FIVE CITIES OF SOUTH SULAWESI PROVINCE", *BAREKENG: Jurnal Ilmu Matematika Dan Terapan*, vol. 19, no. 1, Jan. 2025, doi: https://doi.org/10.30598/barekengvol19iss1pp375-384.
[53] H. M. Nasution and H. Cipta, "ANALISIS SPASIAL DAN TEMPORAL DATA KEJADIAN BENCANA BANJIR DENGAN MODEL GENERALIZED SPACE-TIME AUTOREGRESSIVE INTEGRATED MOVING AVERAGE (GSTARIMA) | JURNAL ABSIS: JURNAL PENDIDIKAN MATEMATIKA DAN MATEMATIKA", *Jurnal ABSIS*, vol. 6, no. 1, pp. 810–825, 2023.doi: https://doi.org/10.30606/absis.v6i1.2171
[54] R. N. E. Pratiwi, S. Wahyuningsih, and M. Siringoringo, "MODEL GENERALIZED SPACE TIME AUTOREGRESSIVE INTEGRATED MOVING AVERAGE", *EKSPONENSIAL*, vol. 10, no. 2, Feb. 2020.
[55] N. M. Huda and N. Imro"ah, "DETERMINATION OF THE BEST WEIGHT MATRIX FOR THE GENERALIZED SPACE TIME AUTOREGRESSIVE (GSTAR) MODEL IN THE COVID-19 CASE ON JAVA ISLAND, INDONESIA", *Spatial Statistics*, vol. 54, Apr. 2023, doi: https://doi.org/10.1016/j.spasta.2023.100734.
[56] I. Iľko, V. Peterkova, J. Maniak, and D. Štefánik, "THE IMPACT OF THE NEW YEAR CELEBRATION ON THE AIR-POLLUTION IN SLOVAKIA", *Journal Of Environmental & Earth Sciences*, vol. 6, no. 3, Sep. 2024, doi: https://doi.org/10.30564/jees.v6i3.7091.
[57] P. Monika, B. N. Ruchjana, and A. S. Abdullah, "GSTARI-X-ARCH MODEL WITH DATA MINING APPROACH FOR FORECASTING CLIMATE IN WEST JAVA", *Computation*, vol. 10, no. 12, Dec. 2022, doi: https://doi.org/10.3390/computation10120204.
[58] M. Akbar, B. Ruchjana, and M. Riyadi, "GSTAR-SUR MODELING WITH CALENDAR VARIATIONS AND INTERVENTION TO FORECAST OUTFLOW OF CURRENCIES IN JAVA INDONESIA", Presented At The Journal Of Physics: Conference Series, IOP Publishing, 2018, doi: https://doi.org/10.1088/1742-6596/974/1/012060
[59] N. M. Huda, U. Mukhaiyar, and N. Imro"ah, "AN ITERATIVE PROCEDURE FOR OUTLIER DETECTION IN GSTAR(1;1) MODEL", *BAREKENG: Jurnal Ilmu Matematika Dan Terapan*, vol. 16, no. 3, Sep. 2022, doi: https://doi.org/10.30598/barekengvol16iss3pp975-984.
[60] P. R. Arum, A. R. Indriani, and M. A. Haris, "FORECASTING THE CONSUMER PRICE INDEX WITH GENERALIZED SPACE-TIME AUTOREGRESSIVE SEEMINGLY UNRELATED REGRESSION (GSTAR-SUR): COMPROMISE REGION AND TIME", *BAREKENG: Jurnal Ilmu Matematika Dan Terapan*, vol. 17, no. 2, Jun. 2023, doi: https://doi.org/10.30598/barekengvol17iss2pp1183-1192.