Journal of Economics and Administrative Sciences
2024; 30(143), pp. 455-473

P-ISSN 2518-5764
E-ISSN 2227-703X

Journal of Economics and Administrative Sciences (JEAS)

Available online at http://jeasiq.uobaghdad.edu.iq
DOI: https://doi.org/10.33095/mxsvs771

Using the Conditional Maximum Likelihood function with the genetic algorithm to estimate STARIMA models (p_λ, d, q_m)

Ali Tareq Abdul Majeed *
Department of Statistics
College of Administration and Economics,
University of Baghdad
Iraq, Baghdad
Ali.abd2101p@coadec.uobaghdad.edu.iq
https://orcid.org/0009-0004-4563-0839
***Corresponding author**

Firas A. Mohammed ALmohana
Department of Statistics
College of Administration and Economics,
University of Baghdad
Iraq, Baghdad
Firas Mohana@coadec.uobaghdad.edu.iq
https://orcid.org/0000-0003-0650-0738

Received:22/1/2024 Accepted:7/4/2024 Published Online First: 1 /10/ 2024

Abstract:

Strategic planning for a specific phenomenon depends mainly on accurate prediction by developing a model to represent that phenomenon. Therefore, this research dealt with spatiotemporal series models such as the STARIMA model (p_λ, d, q_m) model when there is a spatial correlation for neighbouring sites and a spatial correlation for the same geographical location. The methods for estimating the parameters are the Conditional Maximum Likelihood method and the method of estimating the parameters using the genetic algorithm (MLE_{GA}). These models were compared using the statistical comparison metrics (RMSE). The data represents the daily infections with the COVID-19 epidemic for (10) sectors on the Rusafa side of the city of Baghdad from 29/2/2020 until 25/4/2023; it was found that spatiotemporal series data consider the analysis of spatial relationships between geographic points, which allows for a better understanding of the development of phenomena across time and space. The applied results reached the superiority of the spatiotemporal model STARIMA (1_1,1,3_1) with the presence of spatial correlation of neighbouring sites using the genetic algorithm because it has a lower (RMSE), so it was used to predict the daily infections of the Covid-19 epidemic for (10) sectors on the Rusafa side of the city of Baghdad For the period from 26/2/2020 until 5/5/2023.

Paper type :Research paper.

Keywords: spatiotemporal series; STARIMA; weight matrices; Inverse Distance; Conditional Maximum Likelihood Estimation; Genetic Algorithm.

1.Introduction:

Statistics is one of the modern sciences that has a central role in strategic planning and developing appropriate explanations for all vital economic phenomena. The application of statistical methods has a significant impact on scientific conclusions for the near and distant future, as the last decades of the last century witnessed widespread use of time series models, especially autoregressive models - Moving Average, These are called stationary mixed models ARMA(p,q) and unstationary integrated mixed models ARIMA(p,d,q), which treat phenomena over specific periods. Many specialized studies have shown that some time series show spatial correlation, which led to the emergence of another type known as spatiotemporal series, which is concerned with studying time series observations based on nearby locations. They are called integrated autoregressive models for the spatiotemporal moving averages and are known as spatiotemporal mixed series models. STARIMA(p_λ,d,q_m). The data on daily COVID-19 infection cases was presented as a time series in different locations. It is called spatiotemporal data, which studies the relationship influenced by previous times and proximity to the sites. Therefore, many researchers searched for a new technique for modeling the spatiotemporal series. Using the integrated autoregressive model for spatiotemporal moving average, STARIMA(p_λ,d,q_m), the gain is to benefit from the structure of spatiotemporal dependence so that statistical inference can derive strength from neighboring sites in order to obtain accurate predictions that come close to the true values of the studied phenomenon.

1.1 Literature Review:

Modelling spatiotemporal processes requires statistical models that describe the development in the temporal and spatial dimensions of a single variable or multiple variables. These models appeared in the mid-seventies and have increased significantly in recent years because they are closely linked in technology and the provision of large databases through computers. Cliff and Ord (1975) used the spatiotemporal model, after which many techniques compatible with different needs and data types were developed. Pfeiffer and Deutsch (1980) proposed the STARIMA model, an extension of the ARIMA model. Box and Jenkins (1970) used The univariate ARIMA time series, which has been applied for forecasting and analytical purposes in many environmental and agricultural studies, epidemiology, econometrics, transportation methods, and climate. In this part, we discuss the most important previous studies related to the research topic, including: Phillipe (1980) suggested conducting a three-stage iterative modelling to build an autoregressive model for moving environments integrated over space and time, STARIMA, and used data representing the percentage of farms with tractors that represent the central region of the United States for the period from (1920-1964) The conditional maximum likelihood method (MLE) was used to estimate the model parameters. The results of the STARIMA model were compared with previous models that were previously postponed to suit the descriptive ability and predictability of the models. The results showed the STARIMA model's superiority over the rest of the models. Salvador (2019) proved able to compare the results of the univariate ARIMA model with the STARIMA model, as the STARIMA spatio-temporal models proved beneficial in modelling many time series and that the data represented the daily average of radioactivity in Portugal from the year (2008-2017). They represented 13 geographical locations in Portugal. The study concluded that the STARIMA spatiotemporal models give better predictions than the ARIMA models by having lower statistical measures. Fuad et al.(2021) studied the spatial auto-regressive model for moving Average (STARIMA), which includes a small number of parameters, and the temporal model, which includes several parameters (ARIMA), to study the impact of the spread of COVID-19 and its aftermath. Government measures to close areas and declare a curfew in the Kingdom of Saudi Arabia due to the increase in the number of infections daily. The first COVID-19 infection was recorded on 3/19/2020, with a gradual increase in infections as the Kingdom of Saudi Arabia took several measures to control the spread of the epidemic, especially during the Hajj season.

Moreover, Umrah for the year (1441 AH), as restrictions were imposed inside and outside the Kingdom starting from 23/3/2020, and the CVID-19 data was recorded from May 31 to October 11, 2020, in three regions in Saudi Arabia. The results showed that the STARIMA spatiotemporal model is more reliable than the ARIMA models for daily confirmed cases in the three regions. Van Zoest et al. (2022) developed STARIMA for predicting COVID-19 cases to include positive information in the model based on geographic location, in contrast. Unlike previously used models that lack location-determined. The previously used model lacks localization and displays the expected number of cases. This data represents direct and indirect infection indicators in addition to recording national healthcare calls from June 2020 to July 2021 in Uppsala County, Sweden. The results showed the predictive accuracy of the spatiotemporal model based on infection status. The study aims to help the state provide medical supplies, put steps and restrictions in place to reduce transmission and involve the rest of the surrounding region in controlling the infection. Kumar and Sarkar (2023) used the hybrid model when they combined the autoregressive time series model for spatial moving averages (STARMA) with the autoregressive conditional variance - Lagrange multiplier (ARCH-LM) to model temperature and forecast. It is difficult due to the spatial correlation of time series data and non-linearity in the natural assumption. To confront these difficulties, a hybrid model of the STARMA-GARCH model was proposed using data representing monthly maximum temperatures in the state of Bihar (India) for the period from January 1981 to December 2020, as it can be classified into four agricultural climatic zones, and the model was estimated using the conditional maximum potential method. In order to predict maximum temperatures and to reach the best prediction, we used the positive points of the two models, the STARMA and GARCH models. The results showed that the hybrid model was superior to the rest of the models by having the least prediction error due to spatial information and non-linear patterns in the data set. Sukarna et al. (2023) used a set of spatiotemporal models to model COVID-19 infection cases to estimate and predict daily infection cases on the island of Sulawesi, and the data recorded represented daily infections from April 10, 2020, until May 7, 2021. The weight matrix that was used is the matrix based on the inverse weight based on the centre of the region in which the infections were recorded. The results obtained showed the superiority of the STARMA spatiotemporal model by having the lowest trade-off measures (RMSE, MAPE), so it is used in prediction and enables decision-makers to obtain accurate information to formulate future policies and develop possible plans.

The problem of this research was that spatial time series data takes into account the analysis of spatial relationships between different geographical points, which allows a better understanding of the development of phenomena over time and space. Many applications of time series are affected by the values of the phenomenon by similar values in the geographical locations surrounding it, so a variation resulting from the effect will appear. The location of this phenomenon makes it difficult for us to calculate this effect only in the temporal dimension.

The objective research aims to build a spatiotemporal model capable of considering spatial influence to reach good results and predictions close to actual reality using R programming.

2- Material and Methods

2.1 The STARIMA Model

The STARIMA model is given in equations 1,2,3,4.

$$\phi_{p,\lambda}(B)\nabla^d Z_t = \theta_{q,m}(B)\varepsilon_t \qquad (1)$$

$$\phi_{p,\lambda}(B) = 1 - \sum_{k=l}^{p}\sum_{l=0}^{\lambda k} \emptyset_{kl} w_l B^k \qquad (2)$$

$$\nabla^d = (1 - B)^d \qquad (3)$$

 P-ISSN 2518-5764
E-ISSN 2227-703X

$$\theta_{q,m}(B) = 1 - \sum_{k=l}^{q} \sum_{l=0}^{mk} \theta_{kl}\, w_l B^k \qquad (4)$$

Where:
Z_t: Time series.
$\emptyset_{kl}$: Autoregressive parameter, k=time, I= location.
θ_{kl}: Moving average parameter, k=time, I= location.
w: Weight matrix.
p: Autoregressive order
q: Moving average order
λk: Spatial autoregressive order.
mk: Spatial moving average order.
∇^d : Difference Parameter
ε_t: is uncorrelated white noise for time *t* with zero mean and constant variance σ_ε^2.
$E[\varepsilon_i(t)] = 0$

$$E\left[\varepsilon_i(t)\varepsilon_j(t+s)\right] = \begin{cases} \sigma_\varepsilon^2 & i = j\,, s = 0 \\ 0 & otherwise \end{cases} \qquad (5)$$

2.2 Conditional Maximum Likelihood Estimation:

The Maximum Likelihood Estimation (MLE) method is one of the most widely used and common approaches for estimating single-variable Autoregressive Integrated Moving Average (ARIMA) models and spatiotemporal Autoregressive Integrated Moving Average (STARIMA) models. Assuming we have an autoregressive model for moving averages for a stationary time series, the Maximum Likelihood Estimation for

$$\phi = \left[\emptyset_{10}, \emptyset_{11}, \ldots\ldots, \emptyset_{1\lambda_1}, \ldots\ldots, \emptyset_{p0}, \emptyset_{p1}, \ldots\ldots \emptyset_{p\lambda_p}\right]^{/} \qquad (6)$$

$$\Theta = \left[\theta_{10}, \theta_{11}, \ldots\ldots, \theta_{1\lambda_1}, \ldots\ldots, \theta_{q0}, \theta_{q1}, \ldots\ldots \theta_{q\lambda_q}\right]^{/} \qquad (7)$$

Where:
λ: Represents the spatial order (λ=0,1,2,…),

Relying on the equation that involves the error term, which is usually distributed with a mean of zero and constant variance, the Maximum Likelihood Estimation for the error term is written as follows:

$$f((\varepsilon|\phi,\Theta,\sigma^2) = (2\pi)^{-\frac{TN}{2}} \left|\sigma^2 I_{NT}\right|^{-\frac{1}{2}} \exp\left(-\frac{1}{2\sigma^2}\acute{\varepsilon} I \varepsilon\right) =$$

$$(2\pi)^{-\frac{TN}{2}} (\sigma^2)^{-\frac{TN}{2}} \exp\left(-\frac{S(\phi,\Theta)}{2\sigma^2}\right) \qquad (8)$$

Where the sum of squared errors is given by:

$$S(\phi,\Theta) = \acute{\varepsilon} I \varepsilon = \sum_{i=1}^{N} \sum_{t=1}^{T} \varepsilon_i^2(t) \qquad (9)$$

Where :

$$\varepsilon = [\varepsilon_1(1), \ldots, \varepsilon_1(T), \ldots, \varepsilon_N(1), \ldots, \acute{\varepsilon_N}(T)] \qquad (10)$$

The values of the parameters that maximize the maximum potential function are obtained. This is equivalent to finding the of (ϕ, Θ), are obtained, and, in turn, the sum of squared errors S(ϕ, Θ) is minimized by estimating the least squares for (ϕ,Θ). The errors are calculated from equation (11) as following:

$$\varepsilon(t) = z(t) - \sum_{k=1}^{p} \sum_{l=0}^{\lambda_k} \phi k_l\, W^{(l)}\, z(t-k) + \sum_{k=1}^{q} \sum_{l=0}^{mk} \theta k_l\, W^{(l)}\, \varepsilon(t-k) \qquad (11)$$

P-ISSN 2518-5764
E-ISSN 2227-703X

The ϵt represents the random errors.
Since t = (1,…, T) and the parameter values (ϕ, Θ), and considering that the observation values z and errors ϵ are unknown for the previous period, these values are calculated. For a specific selection of parameters ϕ,Θ and initial values (z_*, ε_*), the cumulative values $\varepsilon(\phi, \Theta|z_*, \varepsilon_*, W)$ are computed. Taking the logarithm of the likelihood function for the error term, it is as follows:

$$lnf((\varepsilon|\phi,\theta,\sigma^2) = -\frac{TN}{2}ln(2\pi) - \frac{TN}{2}ln(\sigma^2) - \frac{S(\phi,\theta)}{2\sigma^2} \qquad (12)$$

For a constant σ^2, the conditional Maximum Likelihood Estimation for (ϕ,θ) represents conditional least squares estimates obtained by finding values of (ϕ,θ) that minimize the conditional sum of squares for the function :

$$s_*(\phi,\theta) = \acute{\varepsilon}\varepsilon = \sum_{i=1}^{N}\sum_{t=1}^{T}\epsilon_i(t)^2 \qquad (13)$$

The unconditional probability is obtained using the conditional probability with the values of the elements (z_*, ε_*). One of the procedures is to represent the elements z_* and ε_*, and it equals the unconditional expectation for all values of (z(t), ε(t)), if t is Less than one.
The unconditional expectation for ϵ∗ equals zero when the model does not include the deterministic part, m = 0, and the total expectation for z∗ elements = 0.
The best approximation is to calculate ε's from ε(p+1) while setting ε's value to zero.
IF $\varepsilon_t, Z_t = 0$, The estimator of the error variance is defined according to the following equation:

$$\widehat{\sigma^2} = \left(\frac{S_*(\hat{\phi},\hat{\theta})}{TN}\right) \qquad (14)$$

$(\hat{\phi},\hat{\theta})$:That minimize $S_*(\hat{\phi},\hat{\theta})$.

2.3 Estimating STARMA parameters using genetic algorithms:

2.3.1 Genetic Algorithm:

This method is used to find solutions based on biological evolution and is classified as one of the evolutionary algorithms. It is built on simulating the natural process from the perspective of Darwin's theory. It serves as a randomized search method to find optimal solutions by achieving the principle of optimality. It utilizes natural biological mechanisms such as genetics, mating, and genetic mutation.

2.3.2 Steps of Methodology:

Evolutionary algorithms are implemented using computer simulation programs by considering the smallest unit in the algorithm, representing chromosomes as individuals in their processes to reach optimal solutions. One of the most important ways to represent chromosomes is (binary) encoding, which uses the numbers (0,1). The (evolutionary) process begins by selecting chromosomes randomly from the initial population, and the selection is repeated from generation to generation. In each generation, the fitness value is calculated using the (fitness function) for all chromosomes and individually to choose the optimal chromosomes. Then, the (crossover) operation is performed between the optimal chromosomes, followed by genetic (mutation) application. These processes are repeated until the optimal solution is reached.
As for the steps of parameter estimation, they are as follows:

1 -The initial population is determined by randomly generating K chromosomes with parameters (ϕ and θ) .

2 -Chromosomes are preserved for the next generation (N_{Keep}) based on the values of the differentiation function, and the sum of the squares of the errors of the residuals (RSS) is used as its value. After that, the population is arranged ascendingly from smallest to highest to obtain chromosomes (ϕ, θ) representing a new individual.

3 -The selection involves choosing two chromosomes to represent the parents. Two selection methods are used:

- Tournament method: Two individuals are randomly selected from the current generation, compared based on the fitness function value, and the fitter one is chosen to represent the new generation.
- Roulette wheel method: The roulette wheel is divided into 100 sectors, and individuals from the current generation are distributed among the sectors based on the fitness function value for each individual .
- The roulette wheel is then rolled randomly, and we wait for it to stop at an indicator representing the selected individual.
 The equation for the roulette wheel is as follows:

$$P_{selection\ i} = \frac{F_i}{\sum_{j=1}^{n} F_i} \qquad (15)$$

Where:
$P_{selection\ I}$: Probability of selecting an individual.
Fi: Fitness value for the individual.
n: Number of individuals in the generation.
4- The mating process is carried out on the selected parents to generate two new individuals, which continue to form the new generation. There are different types of mating:
- Single-point crossover: Only one segment is exchanged between the parents. Data exchange occurs between parents, with the condition of avoiding repetition in the exchanged information.
- Multi-point crossover (two or more cutting points): Two or more segments are exchanged between parents, avoiding repetition in the exchanged information.
- Cut and splice: Data from the first chromosome is cut from a region that differs from the cutting region in the second chromosome, leading to differences in chromosome lengths.
5- Mutation: Chromosomes (ϕ, θ) with the smallest RSS will not undergo mutation due to their elite status
6- The evaluation of the population is completed after the renewal process, such as step 2. The steps from 6 to 3 are repeated to obtain the next generation, and the parameter values (ϕ, θ) with the lowest RSS in the last generation are utilized.

2.4 The Spatial Weight Matrix:

The STARIMA model uses the matrix of spatial relationships to quantitatively measure spatial proximity or the nearest neighbour to the spatial time series. Assuming the presence of N locations in the study area and the possibility of a spatial relationship between any two locations, the total spatial relationships represent (n * n), forming the structure of the connectivity matrix. It can be represented as follows :

$$w = \begin{bmatrix} w_{11} & w_{12} & \cdots & w_{1n} \\ w_{21} & w_{22} & \cdots & w_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ w_{n1} & w_{n2} & \cdots & w_{nn} \end{bmatrix}$$

The spatial adjacency matrix (spatial connections) describes spatial relationships by identifying the nearest location to the phenomenon. It is represented as a connected network and helps reduce parameters by selecting the nearest neighbours. Excluding distant sites or neighbours by assigning zero to them contributes to better predictions of the studied phenomenon.

2.4.1 General Conditions for the Weight Matrix:

The weight matrix must be square with a connected path, i.e.,

$w_{ij}^{(l)} \neq 0$.

The unity matrix w(0) is defined as the matrix of connections for all locations with each other. The matrix $\sum_{j=1}^{N} wij^{(l)} = 1$ if :

$$\begin{cases} w_{ij}^{(l)} = cov_{ij}^{l}(k), & l = 1 \\ w_{ij}^{(l)} = cov_{ij}^{l}(k) \setminus \sum_{i=1}^{N} cov_{ij}^{l}(k), & l > 1 \end{cases} \quad (16)$$

The stability condition for the spatial weights matrix ($wij^{(l)}$) is summarized as being (row-normalized), meaning that the sum of the elements in each row equals one. Hence, the column elements of the matrix ($wij^{(l)}$) represent the impact of region (i) on all other regions. In contrast, the row elements signify the impact on region (i) from all other regions. Therefore, the importance of having the matrix ($wij^{(l)}$) row-normalized is explained by the fact that all other regions' influence on each region is neutral. On the other hand, column-normalization indicates that the impact of each region on all other regions is neutral.

2.4.2 Inverse Distance Weight Matrix (IDW): which relies on geographic proximity using Euclidean distance, and the formula for calculating the distance between two points is as follows:

$$d_{ij} = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2} \quad (17)$$

Where :

X: The longitude coordinate

Y: The latitude coordinate

The spatial weights matrix IDW is determined as follows:

$$w_{ij} = \begin{cases} 0 & if \quad i = j \\ \dfrac{\left(1/d_{ij}\right)}{\sum_{i=1}^{N}\left(1/d_{ij}\right)} & if \quad i \neq j \end{cases} \quad (18)$$

$d_{i,j}$: represents the distance for the longitude and latitude coordinates of regions i and j.

3. Discussion of Results:

In this section, the real-world data representing the daily infection series of the COVID-19 pandemic in 10 sectors on the Al-Rusafa side of Baghdad city will be practically applied. The data covers the period from February 29, 2020, to May 5, 2023, comprising 1117 observations. The first 1107 observations are utilized as a control sample to estimate the appropriate model from February 29, 2020, to April 25, 2023. Meanwhile, the remaining ten observations from April 26, 2023, to May 5, 2023, are used as an out-of-sample test to assess the accuracy of predicting daily COVID-19 infections using the STARIMA(p_{λ},d,q_m) model. The researcher begins the chapter by providing a simple statistical description of the spatial-temporal series observations and the sectors where COVID-19 infections were recorded. This is achieved through a set of statistical measures and graphs, aiming to give a general overview of the nature of spatial-temporal data. In analyzing the thesis data, the researcher relies on statistical software (R 3.5.2). The data represents daily records of a spatial-temporal time series for Covid-19 infections. Observations were recorded from February 29, 2020, to May 5, 2023, in ten sectors on the Rasafa side of Baghdad province. The sectors in the Rasafa side of Baghdad are represented based on their proximity to the epicentre of the COVID-19 outbreak as follows**:**

1- Baladiaat 1 Sector, 2- Alsader Cite Sector, 3- Baghdad Al jdida Sector, 4- Baladiaat 2 Sector, 5- Alrisafa Sector, 6- Nahrwan Sector, 7- Adhamiya Sector, 8- Alsha'ab Sector, 9- Almada'in Sector, 10- Istiklal Sector.

Some descriptive statistics were calculated after taking the first differences when (d=1) for all sectors of the spatiotemporal time series of daily COVID-19 infections, as shown in Table (1).

Table 1: Descriptive Statistics for Spatiotemporal Data for the Period from 2020/2/29 to 2023/4/25 when (d=1)

Sectors	Me	Mean	SD	SK	Kur
Baladiaat 1	0	-0.00181	32.208	-0.462	21.667
Alsader City	0	-0.00090	57.887	0.789	26.296
Baghdad Al jdida	0	0.00000	38.628	0.080	12.434
Baladiaat 2	0	0.00000	28.156	1.596	39.499
Alrisafa	0	0.00090	24.432	-0.064	10.273
Nahrawan	0	0.00000	31.561	0.705	21.503
Adhamiya	0	0.00271	38.252	-0.292	11.096
Alsha'ab	0	0.00000	40.465	0.232	16.796
Almada'in	0	0.00000	38.678	0.298	23.109
Istiklal	0	0.00000	17.619	-0.139	13.240

Where: Me= Median, Mean= Arithmetic Mean, SD= Standard Deviation, Sk= Skewness, Kur= Kurtosis.

From Table Table (1), we observe that the difference between the Arithmetic Mean and the median is minimal and close to zero, which indicates that the distribution is symmetrical, and this is demonstrated by the value of the skewness coefficient being less than (-3,3), meaning that the distribution of the spatiotemporal data follows a normal distribution. And then determining the Orders of the STARIMA(p_λ,d,q_m) Model soThe initial spatial time series model, STARIMA(pλ,d,qm), is identified by determining the model orders using the Spatial-Temporal Autocorrelation Function (STACF) and the Spatial-Temporal Partial Autocorrelation Function (STPACF) as follows: After achieving the stability of the spatial time series, the next step is to identify the spatial time series model by determining the order of the Spatial Autoregressive (STAR) component, denoted as (p_λ), and the order of the Spatial Moving Averages (STMA), denoted as (q_m). Consequently, this allows us to specify the order of the spatial time series model, STARMA(p_λ, q_m). These two orders are estimated based on the Spatial Temporal Autocorrelation Function (STACF) and the Spatial Temporal Partial Autocorrelation Function, as illustrated in Figure (1).

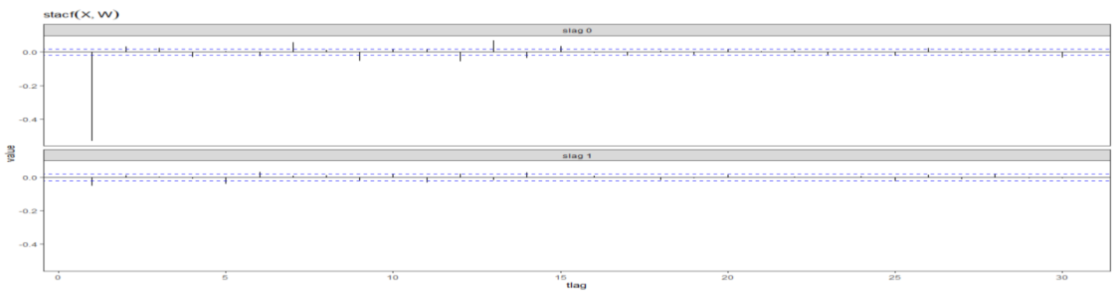

Figure 1: Plot of the Spatial Autocorrelation Function from 2020/2/29 to 2023/4/25.

It can be observed from Figure (1) that the Spatial Autocorrelation Function (STACF) pocket behaviour follows a declining trend With a cutoff, indicating that the model includes a moving average function..

The Partial Spatial Autocorrelation Function (STPACF) is plotted after differencing the spatial time series, as shown in Figure (2).

Figure 2: Plot of the Partial Spatial Autocorrelation Function from 2020/2/29 to 2023/4/25.

We observe oscillations in the cutting process through the Partial Spatial Autocorrelation Function (STPACF), indicating that the model is a mixed STARIMA(p, d, q). The spatial weight matrix was calculated; the onset of the pandemic for the (Baladiaat 1) sector in the Rusafa side of Baghdad Governorate was identified, with the first positive case recorded on 2020/2/29. Subsequently, the infections began to increase, and the nearest vicinity to the (Baladiaat 1) sector was determined based on the calculation of the distance between the sector's center (onset of the pandemic) and the centers of the adjacent sectors. Moreover, the weight matrix (lag=1) depending equation (18), The coordinates of each sector were obtained from the Ministry of Planning - Authority Of Statistics and Geographic Information Systems as follows:

$$\boldsymbol{W} = \begin{bmatrix} 0 & 0.166 & 0.162 & 0.144 & 0.145 & 0.076 & 0.095 & 0.088 & 0.063 & 0.056 \\ 0.145 & 0 & 0.079 & 0.152 & 0.150 & 0.047 & 0.147 & 0.161 & 0.040 & 0.074 \\ 0.201 & 0.112 & 0 & 0.095 & 0.154 & 0.090 & 0.094 & 0.080 & 0.11 & 0.054 \\ 0.168 & 0.203 & 0.089 & 0 & 0.106 & 0.075 & 0.100 & 0.117 & 0.052 & 0.085 \\ 0.140 & 0.166 & 0.120 & 0.088 & 0 & 0.062 & 0.187 & 0.119 & 0.051 & 0.062 \\ 0.148 & 0.105 & 0.142 & 0.125 & 0.099 & 0 & 0.082 & 0.081 & 0.146 & 0.067 \\ 0.090 & 0.160 & 0.072 & 0.082 & 0.184 & 0.040 & 0 & 0.243 & 0.039 & 0.085 \\ 0.085 & 0.178 & 0.062 & 0.097 & 0.119 & 0.040 & 0.248 & 0 & 0.036 & 0.129 \\ 0.129 & 0.095 & 0.190 & 0.093 & 0.109 & 0.154 & 0.086 & 0.078 & 0 & 0.061 \\ 0.092 & 0.139 & 0.072 & 0.120 & 0.105 & 0.056 & 0.147 & 0.217 & 0.049 & 0 \end{bmatrix}$$

The weight matrix (lag=0) calculation for the same vicinity is as follows:

$$W = \begin{array}{cccccccccc} \text{Baladiaat 1} & \text{Alsader City} & \text{Baghdad Al jdida} & \text{Baladiaat 2} & \text{Alrisafa} & \text{Nahrawan} & \text{adhamiya} & \text{Alsha'ab} & \text{Almada'in} & \text{Istiklal} \\ \end{array}$$
$$\begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

Moreover, To estimate parameters for the STARIMA(p_λ, d, q_m) model for the spatiotemporal series of daily Covid-19 infections, we selected the best-fitting model by fitting a set of spatiotemporal series models (STARIMA(p_λ, d, q_m)) based on the available observations and different orders. This was done to choose the optimal model for predicting daily COVID-19 infections, yielding the lowest statistical criteria (AIC, BIC, H-Q). The model selection process was done by writing a program in the R programming language, We observe that the best-extracted models, which have the lowest values for the three statistical criteria (AIC, BIC, H-Q) among the various models When there is a spatial correlation using MLE, are as follows:

1- STARIMA(1_1,1,3_1)

2- STARIMA(1_1,1,5_1).

Moreover, the best-extracted models have the lowest values for the three statistical criteria (AIC, BIC, H-Q). When there is a not spatial correlation using MLE, are as follows:

1- STARIMA(1_0,1,5_0)

2- STARIMA(5_0,1,5_0).

we observe that the best-extracted models, which have the lowest values for the three statistical criteria (AIC, BIC, H-Q) among the various models When there is a spatial correlation using MLE_{GA}, are as follows:

1- STARIMA(1_1,1,1_1)

2- STARIMA(1_1,1,4_1).

Moreover, the best-extracted models have the lowest values for the three statistical criteria (AIC, BIC, H-Q). When there is a not spatial correlation using MLE_{GA}, are as follows:

1- STARIMA(5_0,1,2_0).

After determining the orders of the spatiotemporal model, diagnosed based on the aforementioned statistical criteria, and in order to obtain optimal estimates for the parameters of the spatiotemporal model STARIMA(P_λ, d, Q_m), parameters are estimated using the Conditional Maximum Likelihood Estimation (MLE) and Maximum Likelihood Estimation with Genetic Algorithm (MLE_{GA}) methods, as illustrated in tables (2,3,…and 8).

Table 2: Results of Parameter Estimates for the Spatiotemporal Model (STARIMA(p_λ, d, q_m)) When there is a spatial correlation. Estimated values for the parameters of the STARIMA(1_1,1,3_1) model.

Parameters	MLE	p-value	MLE_{GA}	p-value
φ10	-0.415	<0.001	0.1256	<0.001
φ11	0.1726	<0.001	-0.129	<0.001
θ10	-0.397	<0.001	-0.601	<0.001
θ11	-0.037	<0.001	-0.022	<0.001
θ20	-0.258	<0.001	-0.012	<0.001
θ21	0.254	<0.001	0.118	<0.001
θ30	0.11	<0.001	0.247	<0.001
θ31	0.0267	<0.001	-0.14	<0.001

Table 3: Results of Parameter Estimates for the Spatiotemporal Model (STARIMA(p_λ, d, q_m)) When there is a spatial correlation. Estimated values for the parameters of the STARIMA(1_1,1,5_1) model.

Parameters	MLE	p-value	MLE_{GA}	p-value
φ10	-0.119	0.042	-0.114	0.038
φ11	0.9441	<0.001	-0.58	<0.001
θ10	-0.851	<0.001	-0.192	0.005
θ11	-0.087	0.298	-0.501	<0.001
θ20	0.055	0.380	-0.105	0.044
θ21	-0.028	0.708	0.229	0.002
θ30	0.077	<0.001	0.091	<0.001
θ31	-0.775	<0.001	0.0512	<0.001
θ40	0.8449	<0.001	-0.097	<0.001
θ41	-0.15	<0.001	-0.22	<0.001
θ50	-0.06	<0.001	0.0758	<0.001
θ51	0.0416	0.002	-0.087	<0.001

Table 4: Results of Parameter Estimates for the Spatiotemporal Model (STARIMA(p_λ, d, q_m)) When there is a not spatial correlation. Estimated values for the parameters of the STARIMA(1_0,1,5_0) model.

Parameters	MLE	p-value	MLE_{GA}	p-value
φ10	0.62381	<0.001	-0.27997	<0.001
θ10	-1.53002	<0.001	-0.58664	<0.001
θ20	0.63974	<0.001	-0.10571	0.137
θ30	-0.02470	0.728	-0.08579	<0.001
θ40	-0.07198	<0.001	0.28058	<0.001
θ50	0.07709	<0.001	-0.08953	<0.001

Table 5: Results of Parameter Estimates for the Spatiotemporal Model (STARIMA(p_λ, d, q_m)) When there is a not spatial correlation. Estimated values for the parameters of the STARIMA(5_0,1,5_0) model.

Parameters	MLE	p-value	MLE_{GA}	p-value
φ10	-0.12580	0.133	-0.13649	0.103
φ20	0.19800	0.018	-0.22730	0.007
φ30	0.00748	0.929	0.19735	0.018
φ40	0.19237	0.021	-0.07364	0.369
φ50	-0.03667	0.654	0.35926	<0.001
θ10	-0.78150	<0.001	-0.29461	<0.001
θ20	-0.24375	0.004	-0.02708	0.787
θ30	0.21163	0.030	-0.31073	<0.001
θ40	-0.24058	0.013	0.47187	<0.001
θ50	0.23276	<0.001	-0.09340	0.192

Table 6: Results of Parameter Estimates for the Spatiotemporal Model (STARIMA(p_λ, d, q_m)) When there is a spatial correlation. Estimated values for the parameters of the STARIMA(1_1,1,1_1) model.

Parameters	MLE	p-value	MLE_{GA}	p-value
φ10	-0.101	<0.001	-0.119	<0.001
φ11	0.053	<0.001	-0.171	<0.001
θ10	-0.841	<0.001	-0.669	<0.001
θ11	0.097	<0.001	0.062	<0.001

Table 7: Results of Parameter Estimates for the Spatiotemporal Model (STARIMA(p_λ, d, q_m)) When there is a spatial correlation. Estimated values for the parameters of the STARIMA(1_1,1,4_1) model.

Parameters	MLE	p-value	MLE_{GA}	p-value
φ10	0.1197	0.150	-0.27	<0.001
φ11	0.7093	<0.001	0.2116	0.030
θ10	-1.096	<0.001	-0.229	0.007
θ11	0.168	0.090	-0.102	0.304
θ20	0.001	0.990	-0.055	0.479
θ21	0.036	0.690	-0.023	0.799
θ30	-0.517	<0.001	-0.029	0.007
θ31	0.5931	<0.001	-0.013	0.346
θ40	-0.104	<0.001	-0.103	<0.001
θ41	-0.01	0.443	-0.1	<0.001

Table 8: Results of Parameter Estimates for the Spatiotemporal Model (STARIMA(p_{λ}, d, q_m)) When there is a not spatial correlation. Estimated values for the parameters of the STARIMA(5_0,1,2_0) model.

Parameters	MLE	p-value	MLE_{GA}	p-value
φ10	0.11175	<0.001	-0.04164	<0.001
φ20	-0.12678	<0.001	-0.01988	<0.001
φ30	-0.10848	<0.001	0.11796	<0.001
φ40	-0.12277	<0.001	-0.09494	<0.001
φ50	-0.08215	<0.001	0.01674	<0.001
θ10	-1.01901	<0.001	-0.58388	<0.001
θ20	0.30584	0.170	-0.01262	<0.001

Then, The comparison between the parameter estimation methods for the spatiotemporal model STARIMA(pλ, d, qm) for the (7) models was conducted using the Root Mean Square Error (RMSE) metrics, as illustrated in tables (9,10,…and 15).

Table 9: Results of Comparison metrics for the STARIMA(1_1,1,3_1) model

Parameters	MLE	MLE_{GA}
RMSE	93.42743	52.67869

Table 10: Results of Comparison metrics for the STARIMA(1_1,1,5_1) model

Parameters	MLE	MLE_{GA}
RMSE	159.30736	158.65299

Table 11: Results of Comparison metrics for the STARIMA(1_1,1,1_1) model

Parameters	MLE	MLE_{GA}
RMSE	223.20569	144.66878

Table 12: Results of Comparison metrics for the STARIMA(1_1,1,4_1) model

Parameters	MLE	MLE_{GA}
RMSE	172.88557	128.22305

Table 13: Results of Comparison metrics for the STARIMA(1_0,1,5_0) model

Parameters	MLE	MLE_{GA}
RMSE	186.78538	170.42086

Table 14: Results of Comparison metrics for the STARIMA(5_0,1,5_0) model

Parameters	MLE	MLE_{GA}
RMSE	189.70279	174.84885

Table 15: Results of Comparison metrics for the STARIMA(5_0,1,2_0) model

Parameters	MLE	MLE_{GA}
RMSE	202.67593	180.73876

Through the above tables and the comparison of (140) spatiotemporal models using the best-extracted models, totalling (7) significant models, it is evident from the tables that the best spatiotemporal model is when (P=1, d=1, q=3). All comparison metrics, RMSE, for the methods sequentially (93.42743) with the MLE method and (52.67869) with the MLE_{GA} method represent the smallest values compared to the other models. After confirming that the STARIMA(1_1,1,3_1) model is the best spatiotemporal model for all methods, it becomes evident that the MLE_{GA} method is the most effective compared to the other method since the STARIMA(11,1,31) model represents the best-fit model due to its lower differential measures

(RMSE), residuals were extracted. Autocorrelation functions were calculated, and the Box-Pierce statistic Q was computed, resulting in a value of 43.40905. Upon comparison with the critical value with 60 degrees of freedom at a significance level of 0.05 (which is 79.08194), hypothesis H_0 is accepted.
H_0 = No serial correlation residuals
H_1 = There is serial correlation
The autocorrelation function for the residuals was plotted as shown in Figure (3).

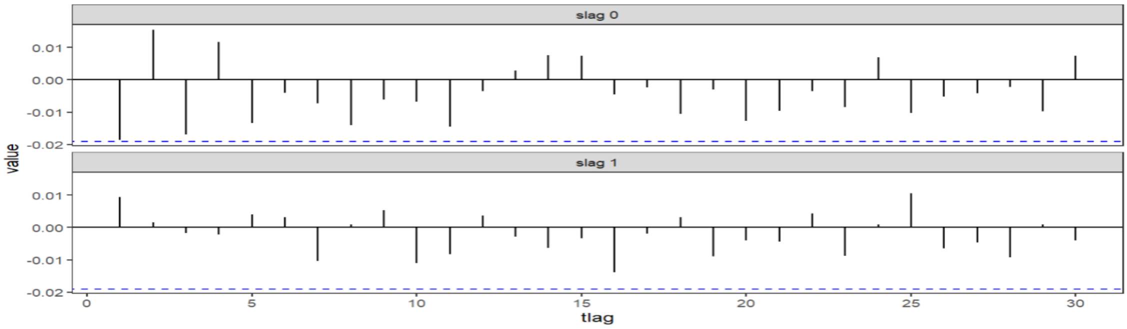

Figure 3: the autocorrelation function plot for the residuals of the spatiotemporal series for the STARIMA(1_1,1,3_1) model.

The figure shows that the extracted residuals are nothing more than uncorrelated random variables. From this, we infer that the model is efficient, good, and suitable for representing this time series. It has been utilized to make future predictions from April 26, 2023, to May 5, 2023.
After diagnosing the model, determining its order, estimating its parameters, and testing it to assess its suitability for representing the spatial-temporal data of COVID-19 infections in the Rusafa sector of Baghdad, it can generate future predictions for the studied phenomenon. The predicted value at t+1 represents the conditional forecast, meaning that:

$$Z^{\wedge(l)}_{t} = E(Z_{t+l}/z_1, z_2, \dots \dots z_n)$$

Predictions were made using the STARIMA(11,1,31) model for ten future values of COVID-19 infections for the period from April 26, 2023, to May 5, 2023, as illustrated in Figures (4,5,…and 13).

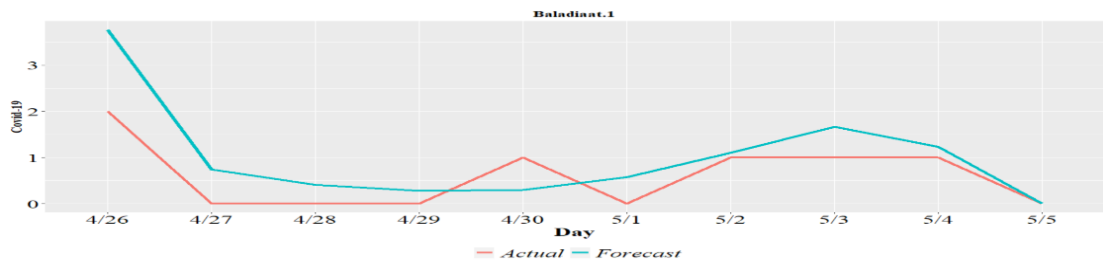

Figure 4: Real and Predicted Values for the STARIMA(1_1,1,3_1) Model from 2023/4/26 to 2023/5/5 for the Baladiaat 1 sector.

Figure 5: Real and Predicted Values for the STARIMA(1_1,1,3_1) Model from 2023/4/26 to 2023/5/5 for the Alsader Cite sector.

P-ISSN 2518-5764
E-ISSN 2227-703X

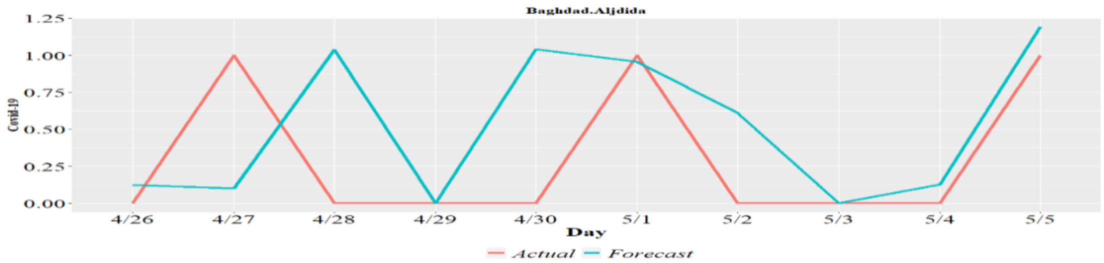

Figure 6: Real and Predicted Values for the STARIMA(1_1,1,3_1) Model from 2023/4/26 to 2023/5/5 for the Baghdad Al jdida sector.

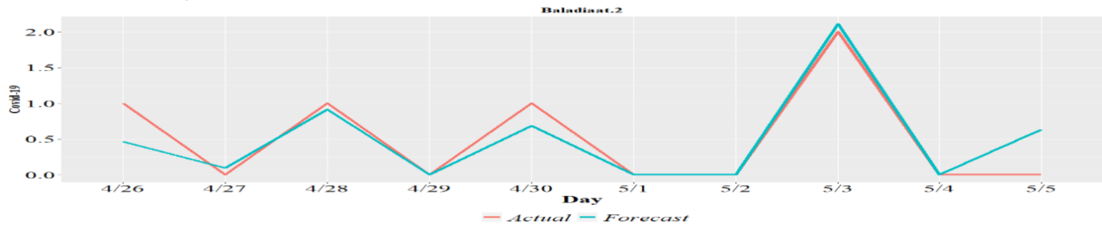

Figure 7: Real and Predicted Values for the STARIMA(1_1,1,3_1) Model from 2023/4/26 to 2023/5/5 for the Baladiaat 2 sector.

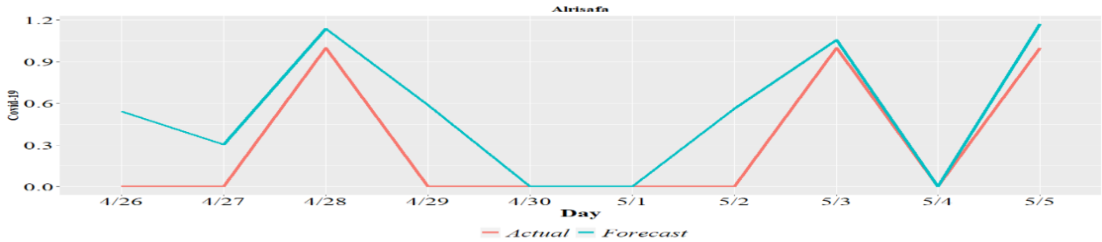

Figure 8: Real and Predicted Values for the STARIMA(1_1,1,3_1) Model from 2023/4/26 to 2023/5/5 for the Alrisafa sector.

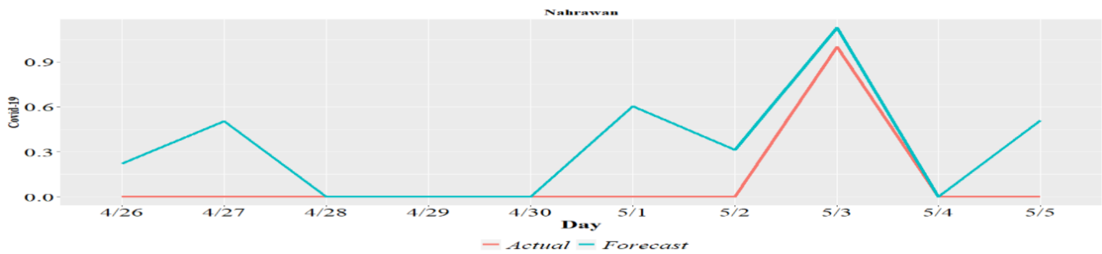

Figure 9: Real and Predicted Values for the STARIMA(1_1,1,3_1) Model from 2023/4/26 to 2023/5/5 for the Nahrwan sector.

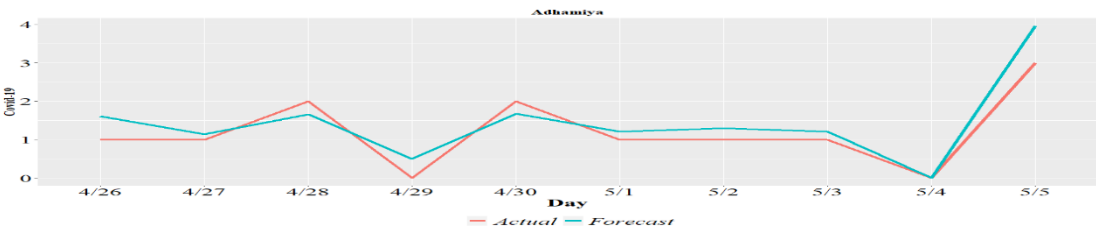

Figure 10: Real and Predicted Values for the STARIMA(1_1,1,3_1) Model from 2023/4/26 to 2023/5/5 for the Adhamiya sector.

P-ISSN 2518-5764
E-ISSN 2227-703X

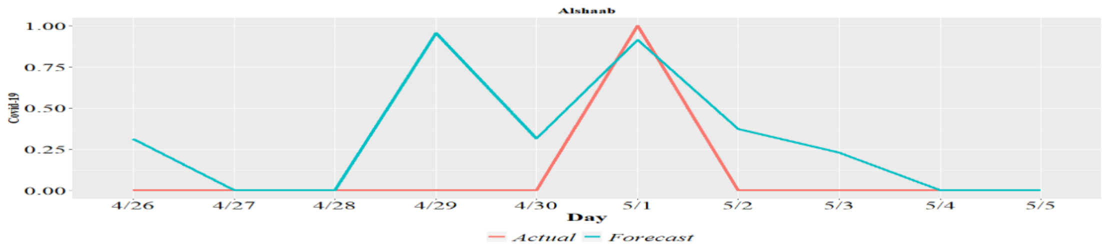

Figure 11: Real and Predicted Values for the STARIMA(1_1,1,3_1) Model from 2023/4/26 to 2023/5/5 for the Alsha'ab sector.

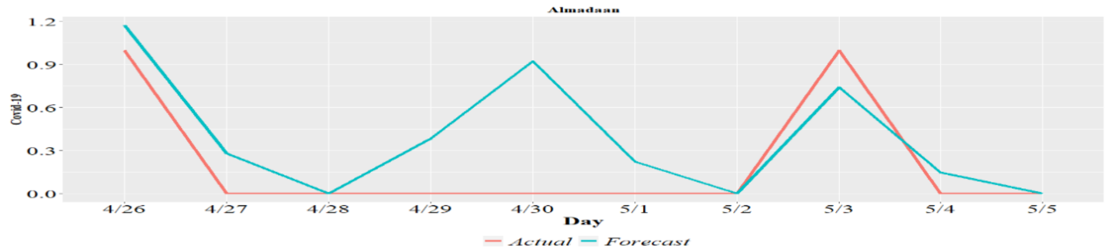

Figure 12: Real and Predicted Values for the STARIMA(1_1,1,3_1) Model from 2023/4/26 to 2023/5/5 for the Almada'in sector.

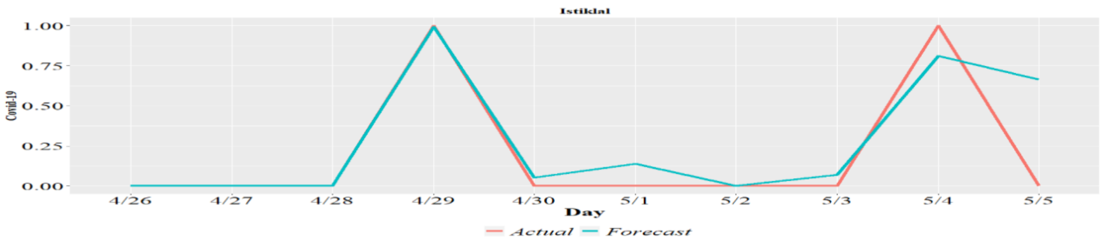

Figure 13: Real and Predicted Values for the STARIMA(1_1,1,3_1) Model from 2023/4/26 to 2023/5/5 for the Istiklal sector.

A comparison between the studied models using forecast accuracy measures was made. Prediction error was determined between the actual and predicted values obtained using the spatiotemporal series model STARIMA(1_1,1,3_1). Accuracy metrics, such as Root Mean Squared Error (RMSE), were utilized to assess the models' prediction effectiveness. Tables (16) illustrate the accuracy metrics for the following models:

Table 16: Accuracy Metrics for STARIMA(1_1,1,3_1) Model

Sector	STARIMA(1_1,1,3_1)	
	MLE	MLE_{GA}
	FRMSE	FRMSE
Baladiaat 1	0.77	0.72
Alsader.City	0.30	0.28
Baghdad.Aljdida	0.63	0.58
Baladiaat 2	0.31	0.28
Alrisafa	0.35	0.33
Nahrawan	0.34	0.32
Adhamiya	0.47	0.44
Alshaab	0.38	0.36
Almadaan	0.39	0.35
Istiklal	0.24	0.22

4. Conclusion:

This applied study, using real data for spatial time series, has produced several conclusions related to the statistical analysis of the individual models (the spatiotemporal model of integrated autoregressive moving Averages, which is symbolized by the symbol STARIMA(pλ, d, qm), the text of which is as follows:

1- The spatial boundaries of the centers of the spread of the epidemic were determined for ten sectors, which are (Municipalities 1, Sadr City, New Baghdad, Municipalities 2, Al-Rusafa, Al-Nahrawan, Al-Adhamiya, Al-Shaab, Al-Mada'in, Al-Istiklal) from the Al-Rusafa side of the city of Baghdad.

2- By studying the beginning of the outbreak of the epidemic in Municipalities Sector 1 and using the weight matrix, the closest neighborhood to the beginning of the spread of the epidemic was determined based on calculating the closest distance between the center of Municipalities Sector 1 and the centers of its neighboring sectors. It was found that the spatiotemporal series model has a smaller number of parameters, which in turn leads to a reduction in the mean square values. It found that the best spatiotemporal predictive model is STARIMA(1_1,1,3_1) for predicting daily cases of COVID-19 infections in the Alrisafa side of Baghdad. The results show Through the results of the table (9)

3- Among the spatiotemporal models, it was found that the most suitable model for the sectors was within the Istiklal sector. The identified model for the time series proved to be the most accurate, obtaining the lowest statistical metrics RMSE (0.22), as shown in Table (16).

4- When comparing estimation methods for the spatiotemporal model STARIMA(1_1,1,3_1), the estimation method using MEl_{GA} showed superiority over the MLE estimation method for the spatiotemporal mode estimation methods for spatiotemporal models. It possesses the lowest accuracy metrics, RMSE, as shown in Table (16).

Authors Declaration:

Conflicts of Interest: None

-We Hereby Confirm That All The Figures and Tables In The Manuscript Are Mine and Ours. Besides, The Figures and Images, Which are Not Mine, Have Been Permitted Republication and Attached to The Manuscript.

- Ethical Clearance: The Research Was Approved By The Local Ethical Committee in The University.

References:

1.Ahmed, A. and Ismael, B. 2022. Using a hybrid SARIMA-NARNN Model to Forecast the Numbers of Infected with (COVID-19) in Iraq. *journal of Economics and Administrative Sciences*. 28(132), pp. 113-118.

2.Alatta, H.J., George, L.E., Behadili, S.F., Sayyid, B.H. and Ali, M.H., 2023. Geospatial Data Analysis of School Distribution in Baghdad City. *Iraqi Journal of Science*, pp.6675-6685.

3.Al-Ramahi, F.K.M., Hasan, M.H. and Zaeen, A.A., 2022. Spatial Analysis of Relative Humidity and Its Effect on Baghdad City for The Years 2008, 2013 and 2018. *Iraqi Journal of Science*, pp.3236-3250.

4.Al-ramahi, F.K.M., Shnain, A.A. and Ali, A.B., 2022. The Modern Techniques in Spatial Analysis to Isolate, Quarantine the Affected Areas and Prevent the Spread of COVID-19 Epidemic. *Iraqi Journal of Science*, pp.4102-4117.

5.Al-ramahi, F.K.M., Shnain, A.A. and Ali, A.B., 2022. The Modern Techniques in Spatial Analysis to Isolate, Quarantine the Affected Areas and Prevent the Spread of COVID-19 Epidemic. *Iraqi Journal of Science*, pp.4102-4117.

6.André, M., Dabo-Niang, S., Soubdhan, T. and Ould-Baba, H., 2016. Predictive spatio-temporal model for spatially sparse global solar radiation data. *Energy*, 111, pp.599-608.

7.Awwad, F.A., Mohamoud, M.A. and Abonazel, M.R., 2021. Estimating COVID-19 cases in Makkah region of Saudi Arabia: Space-time ARIMA modeling. *PLoS One*, 16(4), p.e0250149.
8.Berkani, S., Guermah, B., Zakroum, M. and Ghogho, M., 2023. Spatio-Temporal Forecasting: a Survey of Data-Driven Models using Exogenous Data. *IEEE Access*.
9.Cheng, T., Wang, J., Haworth, J., Heydecker, B. and Chow, A., 2014. A dynamic spatial weight matrix and localized space–time autoregressive integrated moving average for network modeling. *Geographical Analysis*, 46(1), pp.75-97.
10. Cliff, A.D., and Ord, J.K., 1975 Space-Time modelling with an application to regional forecasting. *Transactions of the Institute of British Geographers*, 64, 119-128**.**
11. De Barros, O.M., Marte, C.L., Isler, C.A., Yoshioka, L.R. and da Fonseca Junior, E.S., 2023. Spatial matrices for short-term traffic forecasting based on time series. *Latin American transport studies*, *1*, p.100007.
12. Faisal, F., Novianti, P. and Rizal, J., 2018. The application of spatial analysis and time series in modeling the frequency of earthquake events in bengkulu province. *Aceh International Journal of Science and Technology*, 7(2), pp.103-114.
13. Halim, S., Bisono, I.N., Sunyoto, D. and Gendo, I., 2009, December. Parameter estimation of space-time model using genetic algorithm. In 2009 *IEEE International Conference on Industrial Engineering and Engineering Management (pp. 1371-1375). IEEE*.
14. Hathal, H. and Manfi, S. 2023. An Artificial Intelligence Algorithm to Optimize the Classificationof the HepatitisType. *journal of Economics and Administrative Sciences*. 29(135), pp. 43-55.
15. Kamarianakis, Y. and Prastacos, P., 2003. Space-Time modeling of traffic flow. *The Regional Economics Applications Laboratory*.
16. Kareem, A.M. and Al-Azzawi, S.N., 2021. A stochastic differential equations model for the spread of coronavirus COVID-19): the case of Iraq. *Iraqi Journal of Science*, pp.1025-1035.
17. Khalil, Z. and Ahmmed, F. 2021. Predicting Social Security Fund compensation in Iraq using ARMAX Model. *journal of Economics and Administrative Sciences*. 127(125), pp. 493-508.
18. Kumar, R.R., Sarkar, K.A., Dhakre, D.S. and Bhattacharya, D., 2023. A Hybrid Space–Time Modelling Approach for Forecasting Monthly Temperature. *Environmental Modeling & Assessment*, 28(2), pp.317-330.
19. Munandar, D., Ruchjana, B.N., Abdullah, A.S. and Pardede, H.F., 2023. Literature Review on Integrating Generalized Space-Time Autoregressive Integrated Moving Average (GSTARIMA) and Deep Neural Networks in Machine Learning for Climate Forecasting. *Mathematics*, *11*(13), p.2975.
20. Pfeifer, P.E. and Deutrch, S.J., 1980. A three-stage iterative procedure for space-time modeling phillip. *Technometrics*, 22(1), pp.35-47.
21. Pfeifer, P.E. and Deutsch, S.J., 1980. A STARIMA model-building procedure with application to description and regional forecasting. *Transactions of the Institute of British Geographers*, pp.330-349.
22. Pfeifer, P.E., and Deutsch, S.J., 1981a. Variance of the Sample-Time Autocorrelation Function of Contemporaneously Correlated Variables. *SIAM Journal of Applied Mathematics*, Series A, 40(1), 133-136.
23. Pfeifer, P.E., and Deutsch, S.J., 1981b. Seasonal Space-Time ARIMA modeling. *Geographical Analysis* 13 (2), 117-133.
24. Pfeifer, P.E., and Deutsch, S.J., 1981c. Space-Time ARMA Modeling with contemporaneously correlated innovations. *Technometrics* 23 (4), 410-409.
25. Piantari, E., Rabbani, I.M. and Megasari, R., 2023, October. Seasonal space-time based model for infectious disease prediction. In AIP *Conference Proceedings* (Vol. 2734, No. 1). AIP Publishing.

26. Rathod, S., Gurung, B., Singh, K.N. and Ray, M., 2018. An improved space-time autoregressive moving average (STARMA) model for modelling and forecasting of spatio-temporal time-series data. *Journal of the Indian Society of Agricultural Statistics*, 72(3), pp.239-253.
27. Saad, W. and ALmohana, F. 2022. A Comparison of a Radial Basis Function Neural Network with other Methods for Estimating Missing Values in Univariate Time Series. *journal of Economics and Administrative Sciences*. 28(134), pp. 134-146.
28. Salvador Freire and Rathod, S., 2019. Space Time ARMA Models and their application to Radioactivity in Portugal. *Theoretical and Applied Climatology*, 142, pp.1271-1282.
29. Subba Rao, T. and Costa Antunes, A.M., 2004. Spatio-temporal modelling of temperature time series: a comparative study. In Time Series Analysis and Applications to Geophysical Systems: Part I (pp. 123-150). *Springer New York.*
30. van Zoest, V., Varotsis, G., Menzel, U., Wigren, A., Kennedy, B., Martinell, M. and Fall, T., 2022. Spatio-temporal predictions of COVID-19 test positivity in Uppsala County, Sweden: a comparative approach. *Scientific Reports*, 12(1), p.15176.
31. Zhang, Z., Yan, L. and Gu, Y., 2023, July. ST2T: A Spatio-Temporal Transformer for Cellular Traffic Prediction in Digital Twin Systems. In 2023 *IEEE 6th International Conference on Electronic Information and Communication Technology (ICEICT)* (pp. 1112-1117). IEEE.
32. Zou, J., Zhu, J., Xie, P., Xuan, P. and Lai, X., 2018, August. A STARMA model for wind power space-time series. In 2018 *IEEE Power & Energy Society General Meeting (PESGM)* (pp. 1-5). IEEE.

Journal of Economics and Administrative Sciences
2024; 30(143), pp. 455-473
P-ISSN 2518-5764
E-ISSN 2227-703X

استعمال دالة الإمكان الأعظم الشرطية مع الخوارزمية الجينية في تقدير نماذج $STARIMA(p_{\lambda}, d, q_m)$

علي طارق عبد المجيد[1]
جامعة بغداد/ كلية الإدارة والاقتصاد/ قسم الإحصاء
العراق، بغداد
Ali.abd2101p@coadec.uobaghdad.edu.iq
https://orcid.org/0009-0004-4563-0839

فراس احمد محمد[2]
جامعة بغداد/ كلية الإدارة والاقتصاد/ قسم الإحصاء
العراق، بغداد
Firas Mohana@coadec.uobaghdad.edu.iq
https://orcid.org/0000-0003-0650-0738

Received:22/1/2024 Accepted:7/4/2024 Published Online First: 1 /10/ 2024

مستخلص البحث:

إن التخطيط الاستراتيجي لظاهرة معينة يعتمد بشكل رئيسي على التنبؤ الدقيق من خلال وضع نموذج لتمثيل تلك الظاهرة. لذلك تناول هذا البحث نماذج السلاسل الزمانية المكانية مثل نموذج $STARIMA(p_{\lambda},d,q_m)$ عند عدم وجود ارتباط مكاني للمواقع المجاورة ووجود ارتباط مكاني لنفس الموقع الجغرافي، ان طرق تقدير المعلمات تتمثل بأستخدام طريقة الإمكان الأعظم الشرطية وطريقة تقدير المعلمات باستعمال الخوارزمية الجينية MLE_{GA}. وتمت المقارنة بين هذه النماذج من خلال مقياس المقارنة الاحصائي (RMSE). وان البيانات تمثل الإصابات اليومية بوباء كوفيد 19 لـ (10) قطاعات في جانب الرصافة من مدينة بغداد للمدة من 2020/2/29 لغاية 2023/4/25، إذ وجد ان بيانات السلاسل الزمانية المكانية تأخذ في الاعتبار تحليل العلاقات المكانية بين النقاط الجغرافية مما يسمح بفهم افضل لتطور الظواهر عبر الزمان والمكان. توصلت النتائج التطبيقية الى تفوق الانموذج الزماني المكاني $STARIMA(1_1,1,3_1)$ بوجود الارتباط المكاني للمواقع المجاورة باستخدام الخوارزمية الجينية لإمتلاكه اقل (RMSE) لذا تم استخدامهُ للتنبؤ بالاصابات اليومية لوباء covid-19 لـ (10) قطاعات في جانب الرصافة من مدينة بغداد للفترة من من 2020/2/26 لغاية 2023/5/5.

نوع البحث: ورقة بحثية.

المصطلحات الرئيسة للبحث: السلسلة الزمانية المكانية، انموذج STARIMA، مصفوفة الاوزان، المسافة العكسية، تقدير المعلمات STARIMA ، طريقة الإمكان الأعظم الشرطية، تقدير المعلمات STARIMA بأستعمال الخوارزمية الجينية.

(2)
(3)