Generalized Autocorrelation Function of Stationary Higher Order ARMA Processes: Application to Pandemic Data

Bismark Kwao Nkansah[1*] Pius Gyamu-Atta[1] Henrietta Nkansah[2]

1 Department of Statistics, University of Cape Coast, Ghana

2 Department of Mathematics, University of Cape Coast, Ghana

* E-mail of the corresponding author: bnkansah@ucc.edu.gh

Abstract

The Autocorrelation Function (ACF) of a time series process reveals inherent characteristics of the series that may not be visible from the original series. The ACF of the ARMA(p, q) process has been presented in a few studies in understandably rigorous and laborious manner with no explicit form of the function. In this study, the approach of autocovariance generating functions (acvgf) is used to obtain an explicit expression for a series that follows a linear process under condition of distinct real roots of the AR(p) lag operator polynomial. The technique is used to derive ACFs of processes as far as ARMA($2, q$) for any value of q and subsequently states results for specific ARMA($3, q$) processes. The procedure has shown a clear connection among autocovariances at consecutive lags of the respective process as well as among consecutive orders of the process at particular lags. The derived approach which is applied to daily new Covid-19 cases for countries with stationary series obtains the same results of damp exponential decay in each case as that based on "ARIMAfit" function in R. The results provide useful relations that may be utilized as diagnostic tests for determining whether a given data follows a specified linear process.

Keywords: Autocovariance generating function, linear process, theoretical autocorrelation

1. Introduction

In the domain of stationary time series modelling, the auto-covariance function (acvf) through its associated autocorrelation function (ACF) provides an appealing description of the dependent structure existing between adjacent data points (Carcea and Serfling, 2015; Diebold et al., 2006). ACF is a mathematical representation of the degree of similarity between a given time series and a lagged version of itself over successive time intervals (Khan et al., 2021). By definition, the ACF of a time series process with autocovariance $\gamma(k)$ at lag k is given by

$$\rho(k) = \frac{\gamma(k)}{\gamma(0)} \tag{1}$$

The ACF can either be obtained empirically from the sample data or theoretically from the parameter values based on the appropriate model that characterizes the series. In practice, the sample ACF relates directly to the classical correlation. Literature presents three main approaches on obtaining the theoretical ACF of a time series process: the Yule-Walker approach, comparison of autoregressive and moving average weights, and the use of auto-covariance generating function.

The ACF is a vital tool in time series analysis that reveals hidden characteristics of the underlying data (Mestre et al., 2021; Pardo and Pardo, 2020). In view of this, the literature abounds with the computation of ACF for certain stationary time series processes. Precisely, ACFs are derived (Box et al., 1970; McLeod et al., 1975; Muth, 1978; Triacca, 2016) for lower orders of stationary ARMA models, and a few have focused on higher orders. Presentations of higher order ACFs have been complex and particularly challenging to generalize. The work of McLeod et al. (1975) presents a method for finding the theoretical autocovarince function of an ARMA model. The derivations helps in obtaining an algorithm suitable for machine computation of the theoretical ACF. Although the approach has an advantage of computational simplicity, it does not present the exact analytical expressions of the ACFs. Muth (1978) presents a study on the autocovariance function determined via what is referred to as the z-transform. The autocovariance function's bilateral z-transforms are then generated from the transfer function and inverted after a partial fraction expansion. The results from the approach were used to obtain the autocovariances of certain ARMA(p, q) process. One major drawback of this method is that for partial fraction of cases where the degree of the numerator is higher than the degree of the denominator, the autocovariances are quite arduous to obtain. Specifically, the autocovariances are obtained for ARMA(p, q) processes where $1 \leq p \leq 2$ and $1 \leq q \leq 3$, and no generalizations are made. This article attempts to address these issues by proposing the use of the auto-covariance generating function (acvgf). The approach is presented for ARMA(p, q), for $p = 1, 2,\ \forall\, q$, and some specific processes for $p = 3$. It will be possible therefore to obtain the ACF for the lower order ARMA processes when the expression for the general ARMA(p, q) is known. The derived ACF is subsequently used to approximate the characteristics of relevant pandemic data around some parts of the globe with cases that follow a linear process.

In the next section, we present the underlying methodology for the application of acvgf. Section 3 then provides the results of the derivations of formulas for the linear processes up to ARMA($3, q$). In Section 4, the results are applied to relevant data. The remaining sections discuss pertinent observations and draw conclusion and recommendation.

2. Methods

For a stationary time series process $\{X_t\}$, the sequence of autocovariances $\{\gamma(k)\}$, for $k = 0, 1, \cdots$ can be calculated through a scalar-valued autocovariance generating function (acvgf) defined as

$$c(s) = \sum_{k=-\infty}^{\infty} \gamma(k) s^k \tag{2}$$

This implies that the variance of the process, $\gamma(0)$, is the coefficient of $s^0 = 1$, while $\gamma(k)$ is the coefficient of s^k.

Consider a causal ARMA(p, q) time series process given as

$$\Phi(L)X_t = \Theta(L)Z_t \tag{3}$$

where $\Phi(L)$ and $\Theta(L)$ are linear filters given as

$$\begin{aligned}\Phi(L) =& 1 - \phi_1 L - \phi_2 L^2 - \cdots - \phi_p L^p \\ \Theta(L) =& 1 + \theta_1 L + \theta_2 L^2 + \cdots + \theta_q L^q\end{aligned}$$

and L is the lag operator. Equation (3) can be simplified as

$$X_t = \frac{\Theta(L)}{\Phi(L)} Z_t \tag{4}$$

which is the Moving Average representation of the process, and may be written as

$$X_t = \sum_{r=0}^{\infty} \Psi_r Z_{t-r} \tag{5}$$

where Ψ_r are constants and $\sum \Psi_r^2 < \infty$.

The autocovariance at lag k of Equation (5) is obtained as

$$\begin{aligned}\text{cov}\Big(X_t, X_{t+k}\Big) =& \text{E}\Big(X_t \cdot X_{t+k}\Big) \\ =& \text{E}\Big[\sum \Psi_r Z_{t-r} \cdot \sum \Psi_j Z_{t+k-j}\Big] \\ =& \sigma^2 \sum_{r=0}^{\infty} \Psi_r \Psi_{r+k}\end{aligned}$$

Inferring from Equation (5), we consider a case where

$$c(s) = \sum_{r=0}^{\infty} \Psi_r s^r \tag{6}$$

By multiplying Equation (6) by another power series $c(s^{-1})$, then

$$c(s) \cdot c(s^{-1}) = \text{cov}(X_t, X_{t+k})$$

Thus, $\gamma(k)$ is the coefficient of s^k in the expansion of the power series given by

$$
\begin{aligned}
c(s)c(s^{-1}) =& \sigma^2 \frac{\Theta(s)\Theta(s^{-1})}{\Phi(s)\Phi(s^{-1})} \\
=& \sigma^2 \frac{\Big(\theta_0 + \theta_1 s + \cdots + \theta_p s^p\Big)\Big(\theta_0 + \theta_1 s^{-1} + \cdots + \theta_p s^{-p}\Big)}{\Big(\phi_0 - \phi_1 s - \cdots - \phi_p s^p\Big)\Big(\phi_0 - \phi_1 s^{-1} - \cdots - \phi_p s^{-p}\Big)}
\end{aligned}
\tag{7}
$$

with $\theta_0 = \phi_0 = 1$.

In this study, the underlying assumption is that in Equation (4), the AR(p) lag operator polynomial equation $\Phi(L) = 0$ has p distinct real roots. Hence Equation (7) may be expressed as

$$
c(s)c(s^{-1}) = \sigma^2 \frac{\sum_{j=0}^{q} \theta_j s^j \sum_{j=0}^{q} \theta_j s^{-j}}{\prod_{i=1}^{p}(1-\alpha_i s)(1-\alpha_i s^{-1})} \tag{8}
$$

Obtaining an expression for general values of p and q remains the task to be resolved.

3. Results

Based on Equation (8), this section presents explicit expressions for the ACF of various linear processes.

3.1 ACF of ARMA($1, q$) Process

For ARMA(1,q) process given as

$$
X_t - \phi X_{t-1} = \sum_{j=0}^{q} \theta_j Z_{t-j} \tag{9}
$$

which may be written in terns of lag operator as

$$
(1-\phi L)X_t = \sum_{j=0}^{q} \theta_j L^j Z_t,
$$

the acvgf of the process, following Equation (8) for $p = 1$, simplifies as

$$
\begin{aligned}
c(s)c(s^{-1}) &= \sigma^2 \sum_{r=0}^{\infty}(\phi_1 s)^r \sum_{r=0}^{\infty}(\phi_1 s^{-1})^r \sum_{j=0}^{q} \theta_j s^j \sum_{j=0}^{q} \theta_j s^{-j} \\
&= \sigma^2 \sum_{r=0}^{\infty} \phi^{2r} \Big[\sum_{r=0}^{\infty}(\phi_1 s)^r + \sum_{r=1}^{\infty}(\phi s^{-1})^r\Big] \sum_{j=0}^{q} \theta_j s^j \sum_{j=0}^{q} \theta_j s^{-j}
\end{aligned}
\tag{10}
$$

At lag 0, we consider terms in s^0 and obtain the variance of the ARMA($1, q$) process as

$$
\gamma(0) = \sigma^2 \Bigg\{ \sum_{j=0}^{q} \theta_j^2 + 2 \sum_{n=1}^{q} \sum_{j=0}^{q-n} \phi^n \theta_j \theta_{j+n} \Bigg\} \frac{1}{1-\phi^2} \tag{11}
$$

Subsequently, at lag h, for all $1 \leq h \leq (q-1)$,

$$\begin{aligned}\gamma(h) =& \sigma^2 \Bigg\{ \sum_{n=0}^{h-1} \sum_{j=0}^{q-(n+1)} \phi^{h-1-n} \theta_j \theta_{j+(n+1)} + \sum_{n=0}^{q} \sum_{j=0}^{q-n} \phi^{h+n} \theta_j \theta_{j+n} + \\ & \sum_{n=0}^{q-h-1} \sum_{j=0}^{q-(h+n+1)} \phi^{n+1} \theta_j \theta_{j+(h+n+1)} \Bigg\} \frac{1}{1-\phi^2}\end{aligned} \tag{12}$$

At lag q , we obtain

$$\gamma(q) = \Bigg\{ \sum_{n=0}^{q} \sum_{j=0}^{q-n} \phi^{q-n} \theta_j \theta_{j+n} + \sum_{n=0}^{q-1} \sum_{j=0}^{q-(n+1)} \phi^{q+n+1} \theta_j \theta_{j+(n+1)} \Bigg\} \frac{1}{1-\phi^2} \tag{13}$$

Subsequently, at lag $(q+h)$, for $h \geq 1$ considering terms in s^{q+h} gives

$$\begin{aligned}\gamma(q+h) =& \phi^h \Bigg\{ \sum_{n=0}^{q} \sum_{j=0}^{q-n} \phi^{q-n} \theta_j \theta_{j+n} + \sum_{n=0}^{q-1} \sum_{j=0}^{q-(n+1)} \phi^{q+n+1} \theta_j \theta_{j+(n+1)} \Bigg\} \frac{1}{1-\phi^2} \\ =& \phi^h \gamma(q) \qquad \text{for } h \geq 1 \\ =& \phi^{k-q} \gamma(q) \quad \text{for } k \geq q+1\end{aligned} \tag{14}$$

For example, for ARMA(1,2), following Equations (11) to (14),

$$\rho(k) = \begin{cases} 1 & , k=0 \\ \frac{\phi_1(1+\theta_1^2+\theta_2^2)+(\theta_1+\theta_1\theta_2)[1+\phi_1^2]+\theta_2[\phi_1^3+\phi_1]}{1+\theta_1(\theta_1+2\phi_1)+\theta_2(\theta_2+2\phi_1\theta_1+2\phi_1^2)} & , k=1 \\ \frac{\phi_1^2(1+\theta_1^2+\theta_2^2)+(\theta_1+\theta_1\theta_2)[\phi_1^3+\phi_1]+\theta_2[1+\phi_1^4]}{1+\theta_1(\theta_1+2\phi_1)+\theta_2(\theta_2+2\phi_1\theta_1+2\phi_1^2)} & , k=2 \\ \phi^{k-2}\rho(2) & , k \geq 3 \end{cases} \tag{15}$$

and for ARMA(1,3), it can be deduced that

$$\rho(k) = \begin{cases} 1 & , k=0 \\ \frac{\phi_1(1+\theta_1^2+\theta_2^2+\theta_3^2)+(\theta_1+\theta_1\theta_2+\theta_2\theta_3)[1+\phi_1^2]+(\theta_2+\theta_1\theta_3)[\phi_1^3+\phi_1]+\theta_3[\phi_1^4+\phi_1^2]}{1+\theta_1(\theta_1+2\phi_1)+\theta_2(\theta_2+2\phi_1\theta_1+2\phi_1^2)+\theta_3(\theta_3+2\phi_1\theta_2+2\phi_1^2\theta_1+2\phi_1^3)} & , k=1 \\ \frac{\phi_1^2(1+\theta_1^2+\theta_2^2+\theta_3^2)+(\theta_1+\theta_1\theta_2+\theta_2\theta_3)[\phi_1^3+\phi_1]+(\theta_2+\theta_1\theta_3)[1+\phi_1^4]+\theta_3[\phi_1^5+\phi_1]}{1+\theta_1(\theta_1+2\phi_1)+\theta_2(\theta_2+2\phi_1\theta_1+2\phi_1^2)+\theta_3(\theta_3+2\phi_1\theta_2+2\phi_1^2\theta_1+2\phi_1^3)} & , k=2 \\ \frac{\phi_1^3(1+\theta_1^2+\theta_2^2+\theta_3^2)+(\theta_1+\theta_1\theta_2+\theta_2\theta_3)[\phi_1^4+\phi_1^2]+(\theta_2+\theta_1\theta_3)[\phi_1^5+\phi_1]+\theta_3[1+\phi_1^6]}{1+\theta_1(\theta_1+2\phi_1)+\theta_2(\theta_2+2\phi_1\theta_1+2\phi_1^2)+\theta_3(\theta_3+2\phi_1\theta_2+2\phi_1^2\theta_1+2\phi_1^3)} & , k=3 \\ \phi^{k-3}\rho(3) & , k \geq 4 \end{cases} \tag{16}$$

Clearly, $\rho(k)$ of lower order ARMA processes such as that in Equation (15) can be derived from Equation (16). Similarly, for any higher value of q, expressions for $\rho(k)$ may be obtained from Equations (11) to (14).

It then follows that a relation for three consecutive ACF of ARMA$(1,q)$ is given by

$$\rho^2(q+1) = \rho(q) \times \rho(q+2) \tag{17}$$

3.2 ACF of ARMA$(2,q)$ Process

To obtain the generalized expression for the ARMA$(2,q)$ process, we first consider the ARMA(2,0) process and point out the main expression that facilitates the generalization.

For the ARMA (2,0) process which is given by

$$X_t = \phi_1 X_{t-1} + \phi_2 X_{t-2} + Z_t, \tag{18}$$

from Equation (8) and the result in Equation (10), the acvgf simplifies as

$$c(s)c(s^{-1}) = \sigma^2 \sum_{r=0}^{\infty} \alpha^{2r} \sum_{r=0}^{\infty} \beta^{2r} \left[\sum_{r=0}^{\infty} (\alpha s)^r + \sum_{r=1}^{\infty} (\alpha s^{-1})^r\right]\left[\sum_{r=0}^{\infty} (\beta s)^r + \sum_{r=1}^{\infty} (\beta s^{-1})^r\right] \tag{19}$$

Equation (19) subsequently simplifies as

$$c(s)c(s^{-1}) = \sigma^2 \left(\sum_{r=0}^{\infty} \alpha^{2r} \sum_{r=0}^{\infty} \beta^{2r}\right) T(s^r) \tag{20}$$

where $T(s^r)$ are expressions in terms of s^r. For example, by considering $T(s^0)$, the variance $\gamma_{2,0}(0)$ is obtained as

$$\begin{aligned}
\gamma_{2,0}(0) =& \sigma^2 \Big[\sum_{r=0}^{\infty} \alpha^{2r} \sum_{r=0}^{\infty} \beta^{2r}\Big] T(s^0) \\
=& \frac{\sigma^2}{(1-\alpha\beta)\Big[(1+\alpha\beta)^2 - (\alpha+\beta)^2\Big]} \Big(1+\alpha\beta\Big) \\
=& \frac{\sigma^2}{(1+\phi_2)\Big[(1-\phi_2)^2 - \phi_1^2\Big]} \Big(1-\phi_2\Big)
\end{aligned}$$

noting that $(\alpha+\beta)$=ϕ_1 and $\alpha\beta = -\phi_2$.

Similarly, $\gamma(1)$, $\gamma(2)$ and $\gamma(3)$ are obtained in terms of $\gamma_{2,0}(0)$ as

$$\gamma(1) = \left(\frac{\phi_1}{1-\phi_2}\right)\gamma_{2,0}(0), \quad \gamma(2) = \frac{c_{1,20}}{1-\phi_2}\gamma_{2,0}(0) \text{ and } \quad \gamma(3) = \frac{c_{2,20}}{1-\phi_2}\gamma_{2,0}(0)$$

where $c_{1,20} = \phi_1^2 - \phi_2^2 + \phi_2$ and $c_{2,20} = \phi_1 c_{1,20} + \phi_1\phi_2$

For lags greater or equal to 4 and $r = k - 1$, the generalized autocovariance is given as

$$\begin{aligned}\gamma_{2,0}(k) =& \frac{1}{1-\phi_2}\Bigg[\Big(\phi_1^2+\phi_2\Big)\Bigg\{c_{1,20}\sum_{r-3\geq 2s}\binom{r-3-s}{s}\phi_1^{r-3-2s}\phi_2^{s}+\\ &\sum_{r-4\geq 2s}\binom{r-4-s}{s}\phi_1^{r-3-2s}\phi_2^{s+1}\Bigg\}+\\ &c_{1,20}\sum_{r-4\geq 2s}\binom{r-4-s}{s}\phi_1^{r-3-2s}\phi_2^{s+1}+\\ &\sum_{r-5\geq 2s}\binom{r-5-s}{s}\phi_1^{r-3-2s}\phi_2^{s+2}\Bigg]\gamma_{2,0}(0)\end{aligned} \tag{21}$$

Thus $\rho_{2,0}(k)$, the ACF of the ARMA(2,0), is simply the combinatorial expression coefficient of $\gamma_{2,0}(0)$. Generalizing the procedure for ARMA($2, q$) process given by

$$X_t - \phi_1 X_{t-1} - \phi_2 X_{t-2} = \sum_{j=0}^{q}\theta_j Z_{t-j}, \tag{22}$$

the acvgf of the process, from Equation (8) for $p = 2$, simplifies as

$$c(s)c(s^{-1}) = \sigma^2\sum_{r=0}^{\infty}(\alpha s)^r\cdot\sum_{r=0}^{\infty}(\alpha s^{-1})^r\cdot\sum_{r=0}^{\infty}(\beta s)^r\cdot\sum_{r=0}^{\infty}(\beta s^{-1})^r\sum_{j=0}^{q}\theta_j s^j\sum_{j=0}^{q}\theta_j s^{-j} \tag{23}$$

At lag 0, the variance function is obtained as

$$\begin{aligned}\gamma_{2,q}(0) =& \sum_{r=0}^{\infty}\alpha^{2r}\sum_{r=0}^{\infty}\beta^{2r}\sum_{r=0}^{\infty}(\alpha\beta)^r\Bigg\{\Big[(1+\alpha\beta)\Big]\sum_{j=0}^{q}\theta_j^2+2\Big[(\alpha+\beta)\Big]\sum_{j=0}^{q-1}\theta_j\theta_{j+1}+\\ &2\Big[\alpha\beta(1-\alpha\beta)+(\alpha^2+\beta^2)\Big]\sum_{j=0}^{q-2}\theta_j\theta_{j+2}+2\Big[(\alpha^2\beta+\alpha\beta^2)(1-\alpha\beta)+\\ &(\alpha^3+\beta^3)\Big]\sum_{j=0}^{q-3}\theta_j\theta_{j+3}+\cdots+2\Big[(\alpha^{q-2}\beta+\alpha^{q-3}\beta^2+\cdots+\alpha^2\beta^{q-3}+\\ &\alpha\beta^{q-2})(1-\alpha\beta)+(\alpha^{q-1}+\beta^{q-1})\Big]\sum_{j=0}^{1}\theta_j\theta_{j+(q-1)}+2\Big[(\alpha^{q-1}\beta+\\ &\alpha^{q-2}\beta^2+\cdots+\alpha^2\beta^{q-2}+\alpha\beta^{q-1})(1-\alpha\beta)+(\alpha^q+\beta^q)\Big]\sum_{j=0}^{0}\theta_j\theta_{j+q}\Bigg\}\end{aligned}$$

The expression then simplifies as

$$\gamma_{2,q}(0) = \frac{\sigma^2}{(1-\alpha^2)(1-\beta^2)(1-\alpha\beta)} \Bigg\{ (1+\alpha\beta) \sum_{j=0}^{q} \theta_j^2 + \\ 2\sum_{n=1}^{q} \Big[(1-\alpha\beta) \sum_{r=1}^{n-1} \alpha^{n-r}\beta^r + (\alpha^n + \beta^n) \Big] \sum_{j=0}^{q-n} \theta_j \theta_{j+n} \Bigg\} \tag{24}$$

Subsequently, for $1 \leq h \leq q-1$,

$$\gamma(h) = \frac{\sigma^2}{(1-\alpha^2)(1-\beta^2)(1-\alpha\beta)} \Bigg\{ (1+\alpha\beta) \sum_{j=0}^{q-h} \theta_j \theta_{j+h} + \\ \sum_{n=1}^{h-1} \Big[(\alpha^n + \beta^n) + (1-\alpha\beta) \sum_{r=1}^{h-2} \alpha^{h+n-2-r}\beta^r \Big] \sum_{j=0}^{q-n} \theta_j \theta_{j+n} + \\ \sum_{n=0}^{q} \Big[(\alpha^{n+h} + \beta^{n+h}) + (1-\alpha\beta) \sum_{r=1}^{n} \alpha^{n+h-r}\beta^r \Big] \sum_{j=0}^{q-n} \theta_j \theta_{j+n} + \\ \sum_{n=1}^{q-h} \Big[(\alpha^n + \beta^n) + (1-\alpha\beta) \sum_{r=1}^{n-1} \alpha^{n-r}\beta^r \Big] \sum_{j=0}^{q-(n+h)} \theta_j \theta_{j+(n+h)} \Bigg\},$$

Similarly, at lag q,

$$\gamma(q) = \frac{\sigma^2}{(1-\alpha^2)(1-\beta^2)(1-\alpha\beta)} \sum_{n=0}^{q-1} \Bigg\{ \Big[(\alpha^{q-n} + \beta^{q-n}) + \\ (1-\alpha\beta) \sum_{r=1}^{q-(n+1)} \alpha^{q-n-r}\beta^r \Big] \sum_{j=0}^{q-n} \theta_j \theta_{j+n} + (1+\alpha\beta)\theta_q + \\ \Big[(\alpha^{q+(n+1)} + \beta^{q+(n+1)}) + (1-\alpha\beta) \sum_{r=1}^{q+n} \alpha^{q+1+n-r}\beta^r \Big] \sum_{j=0}^{q-(n+1)} \theta_j \theta_{j+n+1} \Bigg\}$$

and for $h \geq 1$,

$$\gamma(q+h) = \frac{\sigma^2}{(1-\alpha^2)(1-\beta^2)(1-\alpha\beta)} \sum_{n=0}^{q-1} \Bigg\{ \Big[(\alpha^{q+h-n} + \beta^{q+h-n}) + \\ (1-\alpha\beta) \sum_{r=1}^{q+h-(n+1)} \alpha^{q+h-n-r}\beta^r \Big] \sum_{j=0}^{q+h-(n+1)} \theta_j \theta_{j+n} + \\ \Big[(\alpha^{q+h(n+1)} + \beta^{q+h(n+1)}) + (1-\alpha\beta) \sum_{r=1}^{q+h+n} \alpha^{q+h+n-r-1}\beta^r \Big] \\ \sum_{j=0}^{q+h-(n+1)} \theta_j \theta_{j+n+1} \Bigg\}$$

The general expressions are in terms of the roots α and β. Explicit expression for the ACF in terms of the parameters of the ARMA(p, q) may be obtained for a specific process, a task which is quite elusive. For the case of ARMA(1,q) process, however, the task is straightforward.

3.3 Illustration

For example, for ARMA(2,3), the variance may be simplified from Equation (24) as

$$
\begin{aligned}
\gamma_{2,3}(0) =& \frac{\sigma^2}{(1+\phi_2)\left[(1-\phi_2)^2-\phi_1^2\right]}\Bigg\{\left[1-\phi_2\right]\sum_{j=0}^{3}\theta_j^2+\left[2\phi_1\right]\sum_{j=0}^{2}\theta_j\theta_{j+1}+ \\
& \left[2(\phi_1^2+\phi_2-\phi_2^2)\right]\sum_{j=0}^{1}\theta_j\theta_{j+2}+2\theta_3\left[\phi_1^3+2\phi_1\phi_2-\phi_1\phi_2^2\right]\Bigg\}
\end{aligned}
\tag{25}
$$

Equation (25) may be expressed as

$$
\gamma_{2,3}(0) = \frac{1}{(1+\phi_2)\left[(1-\phi_2)^2-\phi_1^2\right]}\sigma^2\left\{1-\phi_2+\theta_1\chi+\theta_2\tau+\theta_3\mu\right\}
$$

where

$$
\begin{aligned}
\chi =& 2\phi_1+\theta_1-\theta_1\phi_2 \\
\tau =& 2\phi_1\theta_1+2\phi_1^2+2\phi_2-2\phi_2^2+\theta_2-\theta_2\phi_2 \\
\mu =& 2\phi_1^3+4\phi_1\phi_2-2\phi_1\phi_2^2+\theta_3-\theta_3\phi_2+2\theta_2\phi_1+2\theta_1\phi_1^2+2\theta_1\phi_2-2\theta_1\phi_2^2
\end{aligned}
$$

and χ and τ are encountered in $\gamma_{2,1}(0)$ and $\gamma_{2,2}(0)$, respectively, of lower order MA processes. Subsequently, for an ARMA(2,3) process, the general autocovariance function is given as

$$
\begin{aligned}
\gamma(k) =& \frac{c_{r,23}}{1-\phi_2+\theta_1\chi+\theta_2\tau+\theta_3\mu}\gamma_{2,3}(0) \\
=& \frac{1}{1-\phi_2+\theta_1\chi+\theta_2\tau+\theta_3\mu}\Bigg\{c_{r,22}+\theta_3\varsigma\left[\sum_{r-4\geq 2s}\binom{(r-4-s)}{s}\phi_1^{r-4-2s}\phi_2^{s+1}\right]+ \\
& \theta_3\nu\left[\sum_{r-3\geq 2s}\binom{(r-3-s)}{s}\phi_1^{r-3-2s}\phi_2^{s}\right]\Bigg\}\gamma_{2,3}(0),
\end{aligned}
\tag{26}
$$

for $r \geq 5$, and $r = k-1$

where

$$
\begin{aligned}
\varsigma =&1+\phi_1^6+5\phi_1^4\phi_2+6\phi_1^2\phi_2^2-\phi_1^4\phi_2^2-3\phi_1^2\phi_2^3-\phi_2+\phi_2^3-\phi_2^4+\theta_3\phi_1^3+2\theta_3\phi_1\phi_2-\theta_3\phi_1\phi_2^2+\theta_2\phi_1^4+\\
&\theta_2\phi_1^2+3\theta_2\phi_1^2\phi_2-\theta_2\phi_1^2\phi_2^2+\theta_2\phi_2-\theta_2\phi_2^3+\theta_1\phi_1^5+\theta_1\phi_1+4\theta_1\phi_1^3\phi_2+3\theta_1\phi_1\phi_2^2-\theta_1\phi_1^3\phi_2^2-2\theta_1\phi_1\phi_2^3\\
\nu =&\phi_1^5+\phi_1^3+4\phi_1^3\phi_2-\phi_1^3\phi_2^2+2\phi_1\phi_2^2+2\phi_1\phi_2-2\phi_1\phi_2^3+\theta_1\phi_1^4+3\theta_1\phi_1^2\phi_2+\theta_1\phi_2^2-\theta_1\phi_1^2\phi_2^2-\theta_1\phi_2^3
\end{aligned}
$$

Similarly, for ARMA(2,2), $\gamma_{2,2}(0)$ can be derived from Equation (25) by simply setting $\theta_3 = 0$. Subsequently, the autocovariance is given as

$$
\begin{aligned}
\gamma(k) =&\frac{1}{1-\phi_2+\theta_1\chi+\theta_2\tau}c_{(k-1),22}\gamma_{2,2}(0)\\
=&\frac{1}{1-\phi_2+\theta_1\chi+\theta_2\tau}\Bigg\{c_{r,21}+\theta_2\lambda\Bigg[\sum_{r-3\geq 2s}\binom{r-3-s}{s}\phi_1^{r-3-2s}\phi_2^{s+1}\Bigg]+\\
&\theta_2\kappa\Bigg[\sum_{r-2\geq 2s}\binom{(r-2-s)}{s}\phi_1^{r-2-2s}\phi_2^{s}\Bigg]\Bigg\}\gamma_{2,2}(0),
\end{aligned}
\tag{27}
$$

for $r \geq 4$, noting that $r = k - 1$

$$
\begin{aligned}
\lambda =&1+\phi_1^4+3\phi_1^2\phi_2-\phi_1^2\phi_2^2-\phi_2+\phi_2^2-\phi_2^3+\theta_2\phi_1^2+\theta_2\phi_2-\theta_2\phi_2^2+\theta_1\phi_1^3+\theta_1\phi_1+2\theta_1\phi_1\phi_2-\theta_1\phi_1\phi_2^2\\
\kappa =&\phi_1^5+\phi_1+4\phi_1^3\phi_2+3\phi_1\phi_2^2-\phi_1^3\phi_2^2-2\phi_1\phi_2^3+\theta_2\phi_1^3+2\theta_2\phi_1\phi_2-\theta_2\phi_1\phi_2^2+\theta_1\phi_1^4+\theta_1\phi_1^2+3\theta_1\phi_1^2\phi_2-\\
&\theta_1\phi_1^2\phi_2^2+\theta_1\phi_2-\theta_1\phi_2^3
\end{aligned}
$$

3.4 Variance of ARMA($3, q$) Process

Generalized acvgf expressions for ARMA($3, q$) and much higher order processes can be obtained following the procedure demonstrated so far. However, they are obviously too lengthy to present for now. Notwithstanding, as a way to validate aspects of the preceding results, we state expressions for the variances of ARMA($3, q$) processes for $0 \leq q \leq 3$. Following Equation (8) the variance of an ARMA($3, q$), for $q = 3$ may be given as

$$
\begin{aligned}
\gamma_{3,3}(0) =&\frac{\sigma^2}{\Big[1+\phi_2+\phi_3(\phi_1-\phi_3^2)\Big]\Big[(1-\phi_2)^2-\phi_1^2-\phi_3(2\phi_1+\phi_3)\Big]}\Bigg\{\Big[1-\phi_2-\phi_3(\phi_1+\phi_3)\Big]\sum_{j=0}^{3}\theta_j^2\\
&+2\Big[\phi_1+\phi_2\phi_3\Big]\sum_{j=0}^{2}\theta_j\theta_{j+1}+2\Big[\phi_1^2+\phi_2-\phi_2^2+\phi_1\phi_3\Big]\sum_{j=0}^{1}\theta_j\theta_{j+2}+\\
&2\theta_3\Big[\phi_1^3+2\phi_1\phi_2-\phi_1\phi_2^2+\phi_3(1-\phi_2-\phi_1\phi_3+\phi_1^2+\phi_2^2-\phi_3^2)\Big]\Bigg\}
\end{aligned}
$$

By putting relevant parameters to zero, the variance of lower order process, such as Equation (25) can be obtained.

3.4 Application to Pandemic Data

This section now uses real data to examine the performance of the derived procedure. In order to relate the derived functions with the literature, the results obtained are compared with existing functions. The data used is obtained from the official website of the Johns Hopkins University Center for Systems Science Engineering (JHU CSSE), and covers the Covid-19 cases for Ghana, Nigeria and South Africa for which cases are found to be stationary and are therefore suitable for the implementations of the results. The data points for all three countries range from January 2020 to March 2022.

Table 1 is a summary of the selected processes that are observed to characterize the daily new Covid-19 cases in the selected countries. The theoretical ACFs are obtained from the parameters of the respective models. It can be seen that all parameter values for the various ARMA processes are statistically significant. The MSE values show greater variability in the performance of the model for South Africa than the other two countries.

Table 1: Summary of appropriate ARMA models of the daily new Covid-19 cases in selected countries

Country	Process	Parameter	Coeff.	SE Coeff.	$p-$value	MSE
Ghana	ARMA(1,4)	ϕ_1	0.9560	0.0137	0.000	76023.67
		θ_1	-0.9326	0.0380	0.001	
		θ_2	0.0248	0.0520	0.011	
		θ_3	0.0250	0.0539	0.028	
		θ_4	0.1373	0.0358	0.003	
Nigeria	ARMA(1,2)	ϕ_1	0.9798	0.0077	0.000	80449.78
		θ_1	-0.8707	0.0371	0.003	
		θ_2	0.1318	0.0338	0.012	
South Africa	ARMA(2,2)	ϕ_1	1.2075	0.1044	0.034	5068631
		ϕ_2	-0.2210	0.1024	0.014	
		θ_1	-0.5621	0.1020	0.009	
		θ_2	-0.1051	0.0594	0.004	

Figures 1, 2, and 3 show the time series plot of the daily cases reported in the three countries, together with their empirical (sample) ACFs. The data are subjected to the "ARIMAfit" function in R, and the models shown in Table 1 are selected for the respective countries. Based on the models, the ACFs based on the McLeod algorithm in R, and the theoretical ACFs based on the derived expressions are obtained, along with the empirical ACF. It can be observed that the sample ACFs show a sinusoidal patterns, an indication that the daily new covid-19 cases in the three countries demonstrate clear patterns. It can again be observed that the times between successive waves are not even, showing that the waves are not necessarily periodic. It is evident from the graphs that the autocorrelations die out at larger lags, an indication that in the distant future, incidence of cases would not be influenced significantly by previous cases. The ACFs based on McLeod's algorithm and that obtained from the derived expressions are found to be almost the same. The theoretical ACFs show that daily cases attenuates exponentially and that incidence of cases would eventually die out.

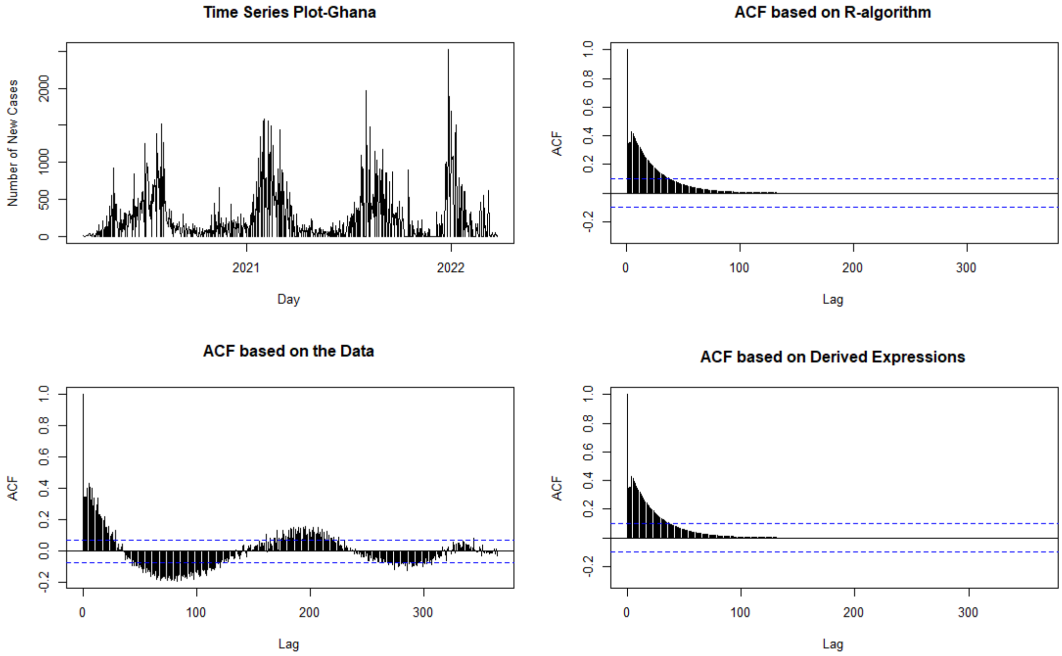

Figure 1: Time series and ACF plots of daily Covid-19 cases for Ghana

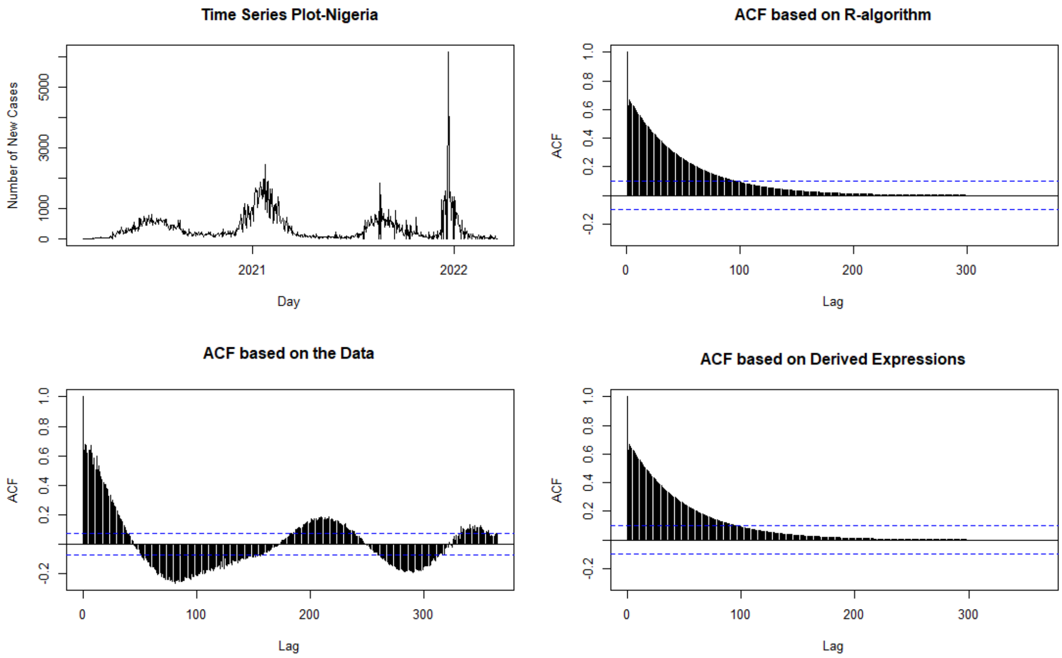

Figure 2: Time series and ACF plots of daily Covid-19 cases for Nigeria

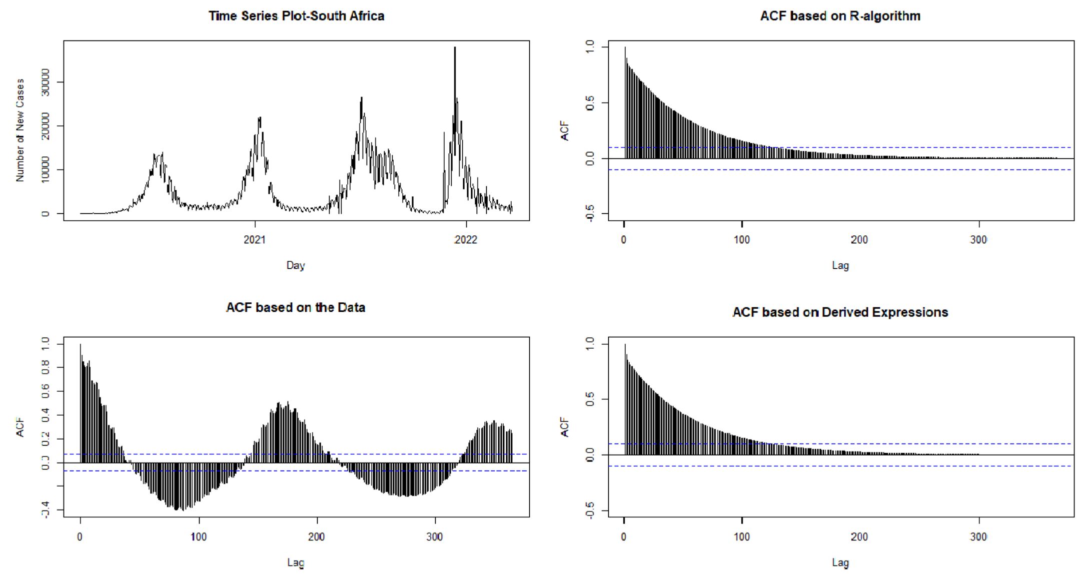

Figure 3: Time series and ACF plots of daily Covid-19 cases for South Africa

4. Discussion

The derivations have shown that the ACF of an ARMA(p, q) process is predominantly influenced by the Moving Average order. In particular, for the ARMA$(1, q)$ and ARMA$(2, q)$ processes, it is seen that ACF at lag q, the order of the MA component, and that beyond lag q are related, and both are unrelated to the ACF preceding lag q. This supports the reason why the ACFs that precede lag q are determined separately.

The literature (Chen et al., 2011; Eshel, 2003) point out the relationship among the autocovariances and autocorrelations obtained from the Yule-Walker (Y-W) simplification. The derivations show that the Y-W recursive formula does not hold for the ACFs at certain lags of some ARMA$(2, q)$ processes, $q \geq 2$. For example, for an ARMA(2,1) process the Y-W recursive formula holds for lag $k \geq 2$, which is consistent with the theory. However, for ARMA(2,2) and (2,3) processes, the Y-W formula holds for lags $k \geq 4$ and $k \geq 5$, respectively.

The slow decay of the theoretical ACFs in Figures 1, 2 and 3 show that the corona virus cases in Ghana, Nigeria and South Africa may be expected to continue for a long time, but will eventually die out. Comparatively, it is expected that daily new Covid-19 cases in Ghana which follows an ARMA(1,4) process with much shorter memory cuts off faster than that of Nigeria and South Africa. This observation agrees with Montgomery et al. (2015) on the theory. Although there is a clear difference between the empirical and theoretical ACFs, both diminish over increasing lags, indicating that incidence of future cases could only be sporadic, and would not follow any discernible pattern. It is also clear that the theoretical ACF is a limiting function of the empirical ACF.

5. Conclusion and Recommendation

The study has presented generalized expressions for the ACFs of ARMA$(1, q)$ and ARMA$(2, q)$ processes for all possible values of q as well as the variances of ARMA$(3, q)$ process for $0 \leq q \leq 3$ under suitable conditions. The generalized ACF helps to establish the connection among consecutive lags and orders of the process. Fast computations are however carried out by iterative implementation of the Y-W equations applied from the appropriately determined lags. The study which is applied to examine the behaviour of the Corona virus pandemic cases in locations where incidence is stationary shows that the pandemic would eventually die out, though there could be sporadic cases not informed by previous cases. The results are in line with non-explicit approach in the literature. It may be relevant to consider ACFs of ARMA(p, q) process under other conditions on the roots of the AR(p) lag operator polynomial.

References

Box, G. E., Jenkins, G. M., and Reinsel, G. (1970). Time series analysis: forecasting and control holden-day san francisco. *BoxTime Series Analysis: Forecasting and Control Holden Day1970*.

Carcea, M. and Serfling, R. (2015). A gini autocovariance function for time series modelling. *Journal of Time Series Analysis*, 36(6):817–838.

Chen, W., Anderson, B. D., Deistler, M., and Filler, A. (2011). Solutions of yule-walker equations for singular ar processes. *Journal of Time Series Analysis*, 32(5):531–538.

Diebold, F. X., Kilian, L., and Nerlove, M. (2006). Time series analysis. In *Macroeconometrics and time series analysis*, pages 317–342. Springer.

Eshel, G. (2003). The yule walker equations for the ar coefficients. *Internet resource*, 2:68–73.

Khan, F., Ali, S., Saeed, A., Kumar, R., and Khan, A. W. (2021). Forecasting daily new infections, deaths and recovery cases due to covid-19 in pakistan by using bayesian dynamic linear models. *PLoS One*, 16(6):e0253367.

McLeod, I. et al. (1975). Derivation of the theoretical autocovariance function of autoregressive-moving average time series. *Applied Statistics*, 24(2):255–256.

Mestre, G., Portela, J., Rice, G., San Roque, A. M., and Alonso, E. (2021). Functional time series model identification and diagnosis by means of auto-and partial autocorrelation analysis. *Computational Statistics & Data Analysis*, 155:107108.

Montgomery, D. C., Jennings, C. L., and Kulahci, M. (2015). *Introduction to time series analysis and forecasting*. John Wiley & Sons.

Muth, E. J. (1978). The autocovariance function determined via the z-transform, with application to box jenkins forecasting models. Technical report, FLORIDA UNIV GAINESVILLE DEPT OF INDUSTRIAL AND SYSTEMS ENGINEERING.

Pardo, S. and Pardo, S. (2020). Autocorrelated data and dynamic systems. *Statistical Analysis of Empirical Data: Methods for Applied Sciences*, pages 197–207.

Triacca, U. (2016). Measuring the distance between sets of arma models. *Econometrics*, 4(3):32.