

## Research Article

# Zero-Inflated Time Series Modelling of COVID-19 Deaths in Ghana

**Kassim Tawiah** , **Wahab Abdul Iddrisu**, and **Killian Asampana Aosega** 

*Department of Mathematics and Statistics, University of Energy and Natural Resources, Sunyani, Ghana*

Correspondence should be addressed to Kassim Tawiah; [kassim.tawiah@uenr.edu.gh](mailto:kassim.tawiah@uenr.edu.gh)

Received 30 January 2021; Revised 13 March 2021; Accepted 7 April 2021; Published 30 April 2021

Academic Editor: Orish Ebere Orisakwe

Copyright © 2021 Kassim Tawiah et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Discrete count time series data with an excessive number of zeros have warranted the development of zero-inflated time series models to incorporate the inflation of zeros and the overdispersion that comes with it. In this paper, we investigated the characteristics of the trend of daily count of COVID-19 deaths in Ghana using zero-inflated models. We envisaged that the trend of COVID-19 deaths per day in Ghana portrays a general increase from the onset of the pandemic in the country to about day 160 after which there is a general decrease onward. We fitted a zero-inflated Poisson autoregressive model and zero-inflated negative binomial autoregressive model to the data in the partial-likelihood framework. The zero-inflated negative binomial autoregressive model outperformed the zero-inflated Poisson autoregressive model. On the other hand, the dynamic zero-inflated Poisson autoregressive model performed better than the dynamic negative binomial autoregressive model. The predicted new death based on the zero-inflated negative binomial autoregressive model indicated that Ghana's COVID-19 death per day will rise sharply few days after 30<sup>th</sup> November 2020 and drastically fall just as in the observed data.

## 1. Introduction

Ghana confirmed its first two cases of the novel coronavirus disease on 12<sup>th</sup> March 2020 at the Noguchi Memorial Institute for Medical Research (NMIMR) [1]. The two cases were all imported. Since then, the government through the Ministry of Health (MoH), the Ghana Health Service (GHS), and other stakeholders introduced prudent measures to help curb the spread of the virus [2]. Key among them was the introduction of the mandatory quarantine of all travellers arriving at the Kotoka International Airport for testing. There was the implementation of social distancing protocols and the compulsory wearing of face/nose masks. Enhanced contact tracing of infected persons and routine surveillance was also instituted. All borders were subsequently closed to travellers. A partial lockdown of Greater Accra and Greater Kumasi was instituted since they were the hot spots. There was ban on all social gatherings. This led to the closure of all public and private schools, night clubs, and churches. Funerals, weddings, and festivals followed suit. However, private funerals and weddings with a maximum of 25 people

with strict adherence to social distancing protocols were permitted. All public transport operators were mandated to reduce their passenger intake in line with the social distancing protocols. The President of Ghana signed an Executive Instrument (EI) to back the ban on social gatherings after Parliament of Ghana passed the Imposition of Restrictions Bill on 21<sup>st</sup> March 2020. There was also the compulsory washing of hands with soap under running water and use of hand sanitizers. Soap and tissue were placed beside Veronica buckets containing water at vantage points throughout the country to enable people wash their hands frequently with soap under running water after every transaction and engagement. Hugging and handshaking were discouraged. The Government introduced stimulus packages for all frontline health workers to boost their efforts in the fight against the pandemic. The Government also cushioned the entire population with free water since water is key to the fight against the spread of SARS-CoV-2. There was fifty percent electricity subsidy for all consumers. Lifeline consumers of electricity were given hundred percent subsidies. These freebies were initially for a period of 1<sup>st</sup>

April 2020 to 30<sup>th</sup> June 2020 but were extended for another 3 months [3, 4].

The partial lockdown of Greater Accra and Greater Kumasi was lifted after almost a month or so in operation. Subsequently, social gathering protocols were relaxed to hundred individuals. All public and private schools were reopened to final years to enable them write their final-year examinations. Restrictions on public transport were also lifted. Restrictions on internal public and private transports on land, sea, and air were lifted. Ghana subsequently opened its international airport to foreign travels from 1<sup>st</sup> September 2020, with strict testing and quarantine rules. Testing results for SARS-CoV-2, the virus that causes COVID-19, 3 weeks prior to arrival at the airport and subsequent testing on arrival were put in place.

Even though the government is making frantic efforts to curb the spread of the coronavirus among the population, the infection figures continue to soar. This calls for an investigation into existing measures and protocols so as to assess their true impact on curtailing the spread of the SARS-CoV-2, the virus that causes COVID-19. Inasmuch as the infection figures are rising, recovery from COVID-19 in Ghana is very impressive.

Maleki et al. [5] purported that for COVID-19 data set, error distribution can cogitate about a two-piece scale mixture of normal (TP-SMN) and designed time series models that work better than ordinary Gaussian and symmetry models. Three regression models (i.e., linear, logarithmic, and quadratic) were proposed for COVID-19 deaths in Pakistan. Influenced by the phase reached by COVID-19 deaths and criteria for assessing goodness of fit, the quadratic model was selected as the best for modelling and predicting death cases in Pakistan [6]. Sperrin and McMillan [4] developed the QCOVID model to predict the risk of COVID-19-related mortality, while the Institute for Health Metrics Evaluation (IHME) proposed and applied deterministic susceptible, exposed, infectious, and recovered (SEIR) compartment frame work model for cases in the United States of America [7].

Dwomoh et al. [8] used mathematical models to investigate COVID-19 infection dynamics in Ghana and delivered a brief forecast of the pandemic trajectory in the country using generalized growth models. They investigated the effective basic reproduction number of the virus in real time applying different techniques of estimation, thereby predicting worse case scenarios amidst integrated individual and Government interventions by the use of compartmental models. Their result indicated that improved individual-level intervention and intensified media coverage can substantially suppress COVID-19 transmission in Ghana and as a result reduce the COVID-19 death rates in the country. However, there seem to be a rise in the daily infection amidst increased Government and media coverage with reduced individual-level intervention. This rise in infections has increased the daily COVID-19-related deaths.

With COVID-19 data from Ghana and Egypt, Asamoah et al. [9] applied sensitivity analysis to suggest that increased diagnoses, enhanced contact tracing, and stringent safety protocols in hospitals or isolation centers with constant

supply of PPEs will help reduce (or possibly stop) the spread of the virus in the two countries.

Bonful et al. [3] audited forty-five public transport stations in the Greater Accra region of Ghana to assess the compliance with the World Health Organization (WHO) safety protocols on the prevention of the spread of COVID-19. These included hand hygiene assessment scale, the availability and use of hand washing facilities, social distancing, and on-going public education on COVID-19 prevention measures. Their findings revealed inadequate washing places, lack of public education on practicing personal hygiene, inadequate alcohol-based sanitizers, and improper face-mask wearing (or no face-mask wearing). They concluded that there is a challenge with COVID-19 prevention compliance.

In this paper, we investigated the current trend of COVID-19 mortalities and use it to predict possible future COVID-19 death trajectory so as to help policy makers to readjust their interventions and strategies. The findings will also help individuals improve their quest to fight the spread of the virus to reduce the number of deaths related to the pandemic. It will also push the media to intensity their coverage to create awareness on what the future holds in relation to possible human life's likely to be lost to COVID-19.

We explored zero-inflated time series models [10–15] to Ghana's COVID-19 death counts per update. The foundation for modelling count data with repeated zeros and overdispersion was provided by Lambert [16], Lee et al. [17], Laird and Ware [18], Min and Agresti [19], Ridout et al. [20], and Yau et al. [21]. Zhu [22] proposed zero-inflated Poisson and negative binomial integer-valued generalized autoregressive conditional heteroscedastic (INGARCH) models, while Yang [14] and Yang et al. [15] proposed a zero-inflated Poisson and negative binomial autoregressive models for zero-inflated and overdispersed discrete count time series data.

We employed zero-inflated time series models as proposed by Yang [14] and Yang et al. [15] as it was best suited for our data. The befitting model obtained was used to predict COVID-19 deaths in Ghana in order to assist the Government and public health experts who are managing the pandemic to know what to expect in terms of deaths before they occur so as to plan ahead.

## 2. Materials and Methods

*2.1. Materials.* We used the entire Ghana as the study setting. Figure 1 presents the map of the study setting showing the number of active and cumulative confirmed COVID-19 cases as at 22<sup>nd</sup> November 2020 in each of the 16 administrative regions of Ghana.

The data for this study consist of confirmed COVID-19 deaths per day from 13<sup>th</sup> March 2020 to 22<sup>nd</sup> November 2020. The daily COVID-19-related deaths were obtained from Our World in Data, an official website for all COVID-19 (<https://ourworldindata.org/coronavirus-source-data>).

*2.2. Methods.* The zero-inflated Poisson (ZIP) and zero-inflated negative binomial (ZINB) autoregressive models

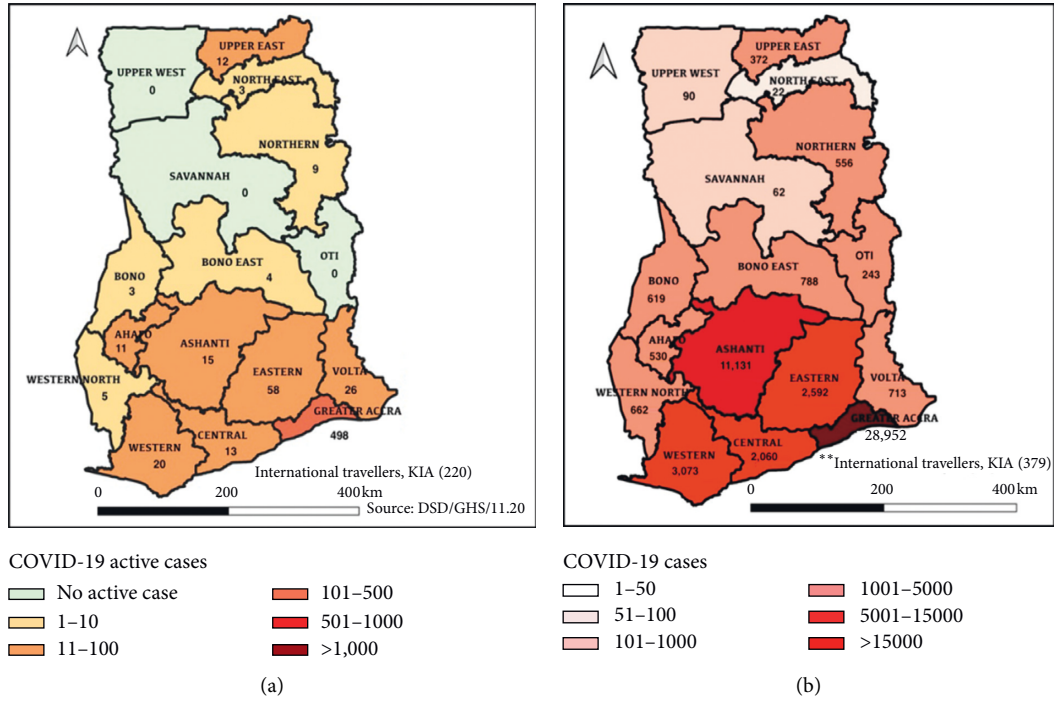


FIGURE 1: Map of study setting (source: <https://www.ghanahealthservice.org/covid19/archive.php#>).

proposed by Yang [14] and Yang et al. [15] were adapted to characterize the trend of COVID-19 deaths in Ghana, which is a discrete count time series with excess zeros.

Let  $\{Y_t\}$  denote the observed COVID-19 deaths, composed of discrete count data which is conditionally distributed as ZIP  $(\lambda_t, \omega_t)$ , where  $\lambda_t$  is the intensity parameter of the baseline Poisson distribution and  $\omega_t$  is the zero-inflation parameter. The zero-inflated Poisson autoregressive (ZIPA) has a probability distribution given by

$$P(Y_t|y_{t-1} = j) = \begin{cases} \omega_t + (1 - \omega_t)\exp(-\lambda_t), & \text{if } j = 0, \\ (1 - \omega_t)\frac{\lambda_t^j \exp(-\lambda_t)}{j!}, & \text{if } j = 1, 2, 3, \dots, \end{cases} \quad (1)$$

where the intensity parameter  $\lambda_t$  and zero-inflation parameter  $\omega_t$  are modelled as follows:

$$\log \lambda_t = X_{t-1}^T \beta, \quad (2)$$

$$\log\left(\frac{\omega_t}{1 - \omega_t}\right) = Z_{t-1}^T \gamma, \quad (3)$$

where  $\beta = (\beta_1, \dots, \beta_p)^T$  and  $\gamma = (\gamma_1, \dots, \gamma_p)^T$  are the regression coefficients for the log-linear part (2) and logistic part (3), respectively. Vectors representing past explanatory

variables which can incorporate functions of the lagged response series accounting for serial correlation are denoted by  $X_{t-1} = (x_{t-1,1}, \dots, x_{t-1,p})^T$  and  $Z_{t-1} = (z_{t-1,1}, \dots, z_{t-1,p})^T$ . The conditional mean and variance of the ZIPA are

$$E(Y_t|y_{t-1}) = \lambda_t(1 - \omega_t), \quad (4)$$

$$\text{Var}(Y_t|y_{t-1}) = (1 - \omega_t)(1 + \lambda_t \omega_t)\lambda_t. \quad (5)$$

Kedem and Fokianos [23] formulated the partial likelihood (PL) of the ZIPA as

$$\text{PL} = \prod_{t=1}^n f_{Y_t}(y_t|y_{t-1}). \quad (6)$$

Substituting (1) into (6) yields

$$\begin{aligned} \text{PL} = & \prod_{t=1}^n I(y_t = 0)[\omega_t + (1 - \omega_t)\exp(-\lambda_t)] \\ & \times \prod_{t=1}^n I(y_t = 1, 2, \dots) \left[ (1 - \omega_t)\frac{\lambda_t^j \exp(-\lambda_t)}{j!} \right]. \end{aligned} \quad (7)$$

Also, when expressions for  $\lambda_t$  in (2) and  $\omega_t$  in (3) are, respectively, put into the above expression, we have

$$\begin{aligned}
 PL &= \prod_{t=1}^n I(y_t = 0) \left[ \frac{\exp(Z_{t-1}^T \gamma)}{1 + \exp(Z_{t-1}^T \gamma)} + \frac{\exp[-\exp(X_{t-1}^T \beta)]}{1 + \exp(Z_{t-1}^T \gamma)} \right] \\
 &\times \prod_{t=1}^n I(y_t = 1, 2, \dots) \left[ \frac{1}{1 + \exp(Z_{t-1}^T \gamma)} \left[ \frac{(\exp(X_{t-1}^T \beta))^j \exp[-\exp(X_{t-1}^T \beta)]}{j!} \right] \right].
 \end{aligned} \tag{8}$$

Yang [14] and Yang et al. [15] developed the Expectation-Maximization (EM) algorithm in the partial likelihood framework to obtain parameter estimates and their standard errors.

Even though the ZIPA may correct for overdispersion [a phenomenon where  $\text{Var}(Y_t|y_{t-1}) > E(Y_t|y_{t-1})$ ] in

discrete count time series data with excess zeros, we extended the ZIPA to zero-inflated negative binomial autoregressive (ZINBA) model which is well known for overdispersed data.

For the ZINBA, the probability distribution is given by

$$P(y_t | y_{t-1} = j) = \begin{cases} \omega_t + (1 - \omega_t) \left( \frac{k_t}{k_t + \lambda_t} \right)^{k_t}, & \text{if } j = 0, \\ (1 - \omega_t) \frac{\Gamma(k_t + y_t)}{\Gamma(k_t) y_t!} \left( \frac{k_t}{k_t + \lambda_t} \right)^{k_t} \left( \frac{\lambda_t}{k_t + \lambda_t} \right)^{y_t}, & \text{if } j = 1, 2, 3, \dots, \end{cases} \tag{9}$$

with  $\lambda_t$  and  $\omega_t$  modelled as in (2) and (3) respectively. The dispersion parameter,  $k_t$ , is modelled as

$$\log k_t = S_{t-1}^T \alpha, \tag{10}$$

where  $\alpha = (\alpha_1, \dots, \alpha_p)^T$  is the regression coefficients and  $S_{t-1} = (s_{t-1,1}, \dots, s_{t-1,p})^T$  is a vector of past explanatory variables.

The mean and variance expressions for the ZINBA are the same as that of the ZIPA stated in (4) and (5), respectively. As in the ZIPA, the PL of the ZINBA is the same as in (6).

Thus, substituting (7) into (6) gives the PL

$$PL = \prod_{t=1}^n I(y_t = 0) \left[ \omega_t + (1 - \omega_t) \left( \frac{k_t}{k_t + \lambda_t} \right)^{k_t} \right] \times \prod_{t=1}^n I(y_t = 1, 2, \dots) \left[ (1 - \omega_t) \frac{\Gamma(k_t + y_t)}{\Gamma(k_t) y_t!} \left( \frac{k_t}{k_t + \lambda_t} \right)^{k_t} \left( \frac{\lambda_t}{k_t + \lambda_t} \right)^{y_t} \right]. \tag{11}$$

When expressions for  $k_t$  in (10),  $\lambda_t$  in (2), and  $\omega_t$  in (3), respectively, are substituted into the PL above, it yields

$$\begin{aligned}
 PL &= \prod_{t=1}^n I(y_t = 0) \left[ \frac{\exp(Z_{t-1}^T \gamma)}{1 + \exp(Z_{t-1}^T \gamma)} + \left( \frac{1}{1 + \exp(Z_{t-1}^T \gamma)} \right) \left( \frac{\exp(S_{t-1}^T \alpha)}{\exp(S_{t-1}^T \alpha) + \exp(X_{t-1}^T \beta)} \right)^{\exp(S_{t-1}^T \alpha)} \right] \\
 &\times \prod_{t=1}^n I(y_t = 1, 2, \dots) \left[ \left( \frac{1}{1 + \exp(Z_{t-1}^T \gamma)} \right) \frac{\Gamma(\exp(S_{t-1}^T \alpha) + y_t)}{\Gamma[\exp(S_{t-1}^T \alpha)] y_t!} \left( \frac{\exp(S_{t-1}^T \alpha)}{\exp(S_{t-1}^T \alpha) + \exp(X_{t-1}^T \beta)} \right)^{\exp(S_{t-1}^T \alpha)} \left( \frac{\exp(X_{t-1}^T \beta)}{\exp(S_{t-1}^T \alpha) + \exp(X_{t-1}^T \beta)} \right)^{y_t} \right].
 \end{aligned} \tag{12}$$

Parameter estimates and their corresponding standard errors in the PL above can be obtained by employing the EM algorithm proposed by Yang [14] and Yang et al. [15].

The models were compared based on the Akaike Information Criterion (AIC; [23, 24]), Bayesian Information Criterion (BIC; [23, 25]), and Takeuchi Information

Criterion (TIC; [26]). These metrics combine a measure of model fit, typically twice the negative log-partial likelihood, with a penalty for model complexity, expressed as a function of the number of parameters [12]. The AIC and BIC are computed by the expressions:

$$\begin{aligned} \text{AIC} &= -2 \log \text{PL} + 2k, \\ \text{BIC} &= -2 \log \text{PL} + k \log(n), \end{aligned} \quad (13)$$

where  $k$  is the number of parameters in the model and  $n$  is the number of observations.

The TIC is calculated by the expression:

$$\text{TIC} = -2 \log \text{PL} + 2\text{tr}\{J_n I_n^{-1}\}, \quad (14)$$

where  $I_n^{-1}$  is the information matrix and  $J_n$  is given by

$$J_n = \sum_{t=1}^n \left\{ \frac{\partial \log f_{Y_t}(y_t | y_{t-1})}{\partial \theta} \right\} \left\{ \frac{\partial \log f_{Y_t}(y_t | y_{t-1})}{\partial \theta} \right\}^T, \quad (15)$$

with  $\theta = (\beta, \gamma)^T$ .

Model fitting was done in  $R$  [27] using the “ZIM” and the `countreg` packages [14, 15, 28].

### 3. Results and Discussion

The trend of COVID-19 death in Ghana as illustrated in Figure 2 gives the impression of a general increase in the death toll from day zero (the day COVID-19 was first discovered in Ghana) to about day 160 after which there is an impression of a general decrease in the death toll to day 250 and beyond. The increase in the number of deaths was expected as the infection rate and active and severe cases continued to soar from day zero to day 150. The decrease in the number of active and severe cases amidst a rise in the infection rate could also be attributed to the decline in the number of deaths. However, a thorough investigation is needed to be carried out on the rise and fall of the deaths to ascertain what truly fuelled it.

As indicated by the histogram (Figure 3.), there is a higher proportion of zero counts (no deaths) per day making up 69.02% of the entire time series data. This is clearly an indication of zero inflation in the data. Even though infection rate continues to rise, the number of deaths being reported in most days ought to be looked into in order to confirm whether the majority of the COVID-19 patients in Ghana have developed resistance to the pandemic or they responded positive to care procedures meted to them at COVID-19 treatment centers.

In order to realize the most apt zero-inflated time series model to characterize the trend of our data, we fitted the ZIPA and ZINBA (Table 1). In the log-linear part of the models, the intercept of the ZIPA was significant at 0.05 significance level, but that of the ZINBA was not. However, the ZINBA has a smaller estimate (0.5821) compared to the estimate (1.1317) of the ZIPA. The `count.lag` and the `trend` were both not significant at 0.05. Fascinatingly, the `count.lag` of the ZIPA model had an increasing effect on the log of the expected number of COVID-19 deaths in Ghana, while it

had a decreasing effect on that of the ZINBA model. The trend was also not significant in both the ZIPA model and the ZINBA model. Also, worth noting is the fact that even though the trend had an increasing effect on the log of expected number of COVID-19 deaths in both models, the effect is higher in the ZINBA than the ZIPA model.

On the logistic part of our models, the intercept of the ZIPA was significant with a higher estimate as compared to the nonsignificant intercept of the ZINBA with a smaller estimate. Just as in the log-linear part, the trend was not significant in the logistic part of the ZIPA and ZINBA models. In both models, the trend has an increasing effect on the log of the expected odds of the number of COVID-19 deaths in Ghana with the odds in the ZINBA model being higher than that of the ZIPA model.

The ZIPA model has higher AIC, BIC, and TIC values than the ZINBA model. For each of these criteria for assessing goodness of fit, there is a tart reduction from the value of the ZIPA to that of the ZINBA. This clearly shows that the ZINBA model may have corrected for more complexity in the data than the ZIPA model.

The test for overdispersion conducted (Table 2) had a score test of 8.3470 and a  $P$  value less than 0.0001. This means that the ZINBA model did well with respect to the overdispersion in our data.

Output from the dynamic zero-inflated Poisson autoregressive (DZIPA) and dynamic zero-inflated negative binomial autoregressive (DZINBA) models, with 200 replications, 100 iterations, and sample size of 200 in each model, are presented in Table 3. The zero-inflation parameter dwindled from the DZIPA (0.6148) to the DZINBA (0.6108). This could mean that the DZIPA may have detected more zeros in the data than the DZINBA. The standard deviation was higher in the DZIPA than in the DZINBA.

In the log-linear part of the models, the intercept is significant in the DZIPA model, but not significant in the DZINBA model. The trend was not significant in both the DZIPA model and the DZINBA model. Nevertheless, the trend had a decreasing effect on the log of the expected number of COVID-19 deaths in Ghana with the DZIPA model having a greater decreasing effect than the DZINBA model. We can deduce that the dynamic models generally forecast a decrease in the expected number of COVID-19 deaths in Ghana.

In respect of the autoregressive part, AR (1) was significant in the DZIPA model as well as the DZINBA model. Thus, we can deduce the DZIPA and the DZINBA models are both AR (1).

There was an unsubstantial increase in the AIC and BIC values from the DZIPA to the DZINBA. The TIC values, however, registered a quantum increase from the DZIPA model to the DZINBA model. Consequently, we can assert that the DZIPA has corrected for more complexity in the data than the DZINBA. With respect to the AIC, BIC, and TIC values, the DZIPA model outperformed the ZIPA model. Notwithstanding, the DZINBA outperformed the ZINBA only in terms of the AIC and BIC values while the opposite is true for the TIC value.

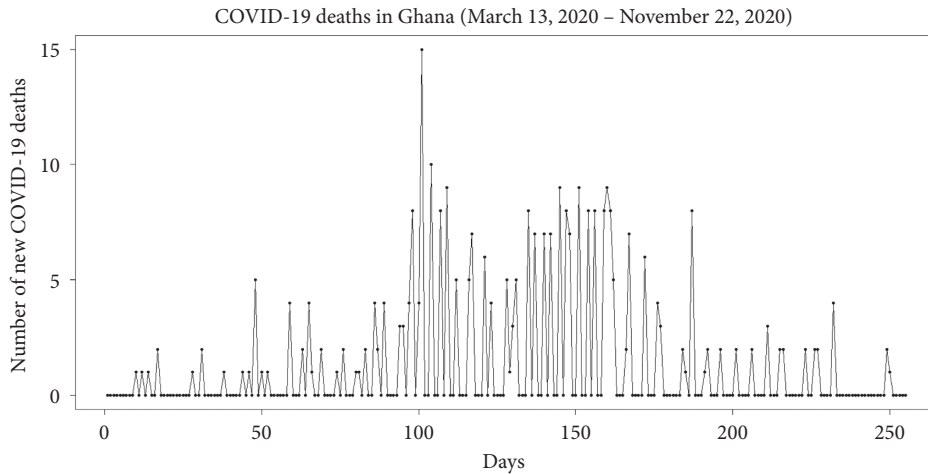


FIGURE 2: Trend of COVID-19 deaths in Ghana from 13<sup>th</sup> March 2020 to 22<sup>nd</sup> November 2020.

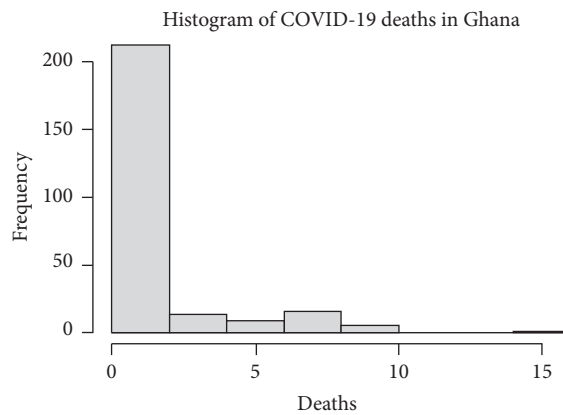


FIGURE 3: Histogram of COVID-19 deaths in Ghana from 13<sup>th</sup> March 2020 to 22<sup>nd</sup> November 2020.

TABLE 1: ZIPA and ZINBA models estimate with corresponding standard errors (s.e.) for COVID-19 deaths in Ghana.

Coefficients	ZIPA		ZINBA	
	Estimate (s.e.)	Pr ( $>  z $ )	Estimate (s.e.)	Pr ( $>  z $ )
Intercept	1.1317 (0.1425)	<0.0001	0.5821 (0.4654)	0.2210
Count.lag1	0.0580 (0.1309)	0.6578	-0.0391 (0.2581)	0.8797
Trend	1.8618 (0.9858)	0.0590	4.9987 (3.2218)	0.1208
	Logistic		Logistic	
Intercept	0.6342 (0.2774)	0.0223	0.0896 (0.5284)	0.8654
Trend	1.0210 (1.8822)	0.5875	3.0297 (2.8942)	0.2952
Dispersion, $k_t$	—	—	1.5353	—
Goodness of fit				
AIC	720.7728	—	681.6503	—
BIC	738.4595	—	702.8743	—
TIC	729.1185	—	683.6137	—

TABLE 2: Test for overdispersion.

	$H_0$ : ZIPA
	$H_1$ : ZINBA
Score.test	8.3470
P value	<0.0001

Trace plots for the DZIPA and DZINBA models are presented in Figure 4. From the plots, we can see that the partial likelihood becomes progressively greater conspicuously preceding all others in time for several iterations and then maintains stability as the estimated parameters become very close to the maximum likelihood estimator [14, 15].

Figure 5 points to a probability integral transform (PIT) histogram [29], which appears to approach uniformity. The horizontal line depicts the count that each of the bins would have if the histogram was perfectly uniform. Hence, the probabilistic calibration of the fitted ZINBA model is sufficient.

Time series of the observed daily new deaths of COVID-19 from 23<sup>rd</sup> November 2020 to 6<sup>th</sup> December 2020 and predicted daily new deaths based on the fitted ZINBA model are shown in Figure 6. It is observed that the overall trend of the two

TABLE 3: Dynamic zero-inflated Poisson autoregressive (DZIPA) model and dynamic negative binomial autoregressive (DZINBA) model estimate with corresponding standard errors (s.e.) for COVID-19 deaths in Ghana.

Coefficients	DZIPA Log-linear		DZINBA Log-linear	
	Estimate (s.e.)	Pr ( $>  z $ )	Estimate	Pr ( $>  z $ )
Intercept	1.1012 (0.5469)	0.0441	0.8066 (1.2224)	0.5093
Trend	-1.3129 (3.9978)	0.7426	-0.3020 (7.4510)	0.9677
Autoregressive				
AR1	0.9735 (0.0172)	<0.0001	0.9618 (0.0291)	<0.0001
Zero-inflation	0.6148	—	0.6108	—
Standard deviation	0.1908	—	0.2285	—
$N$	200	—	200	—
$R$	200	—	200	—
Number of iterations	100	—	100	—
Dispersion, $k_t$	—	—	11.3647	—
Goodness of fit				
AIC	629.1376	—	636.9148	—
BIC	646.8336	—	658.1624	—
TIC	722.0860	—	1023.5070	—

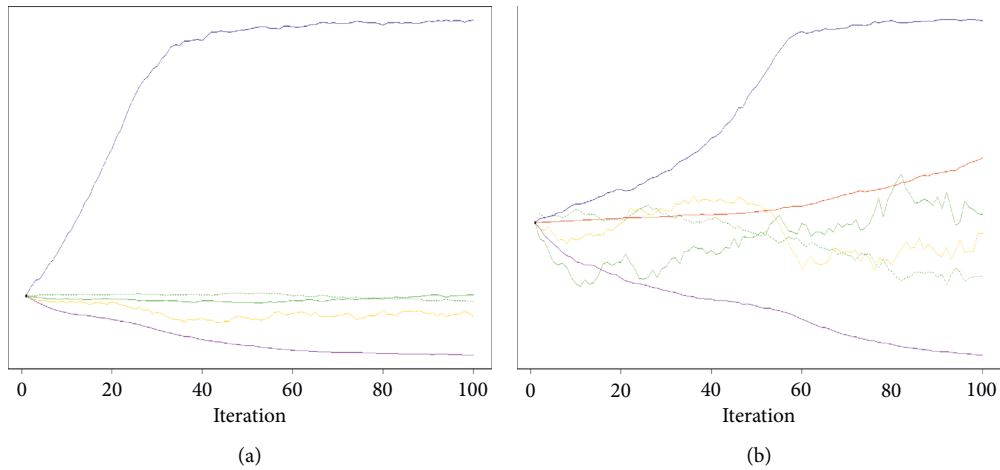


FIGURE 4: Trace plots from the DZIPA model (a) and DZINBA model (b) of COVID-19 deaths in Ghana.

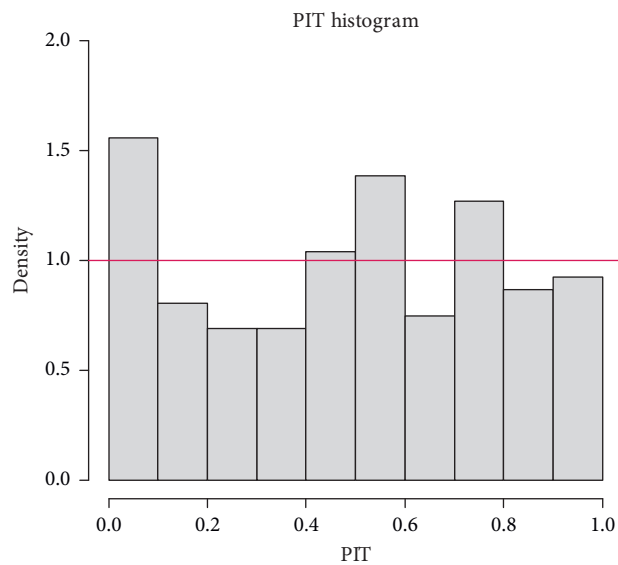


FIGURE 5: Probability integral transform (PIT) histogram for the ZINBA model.

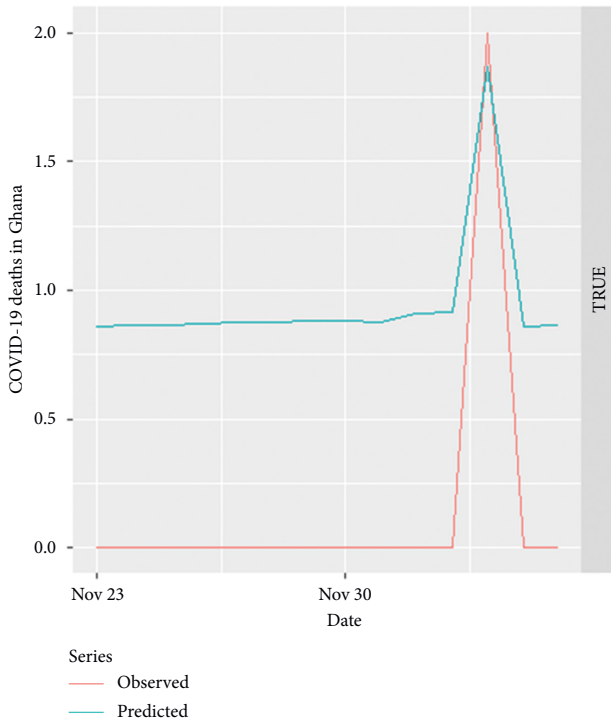


FIGURE 6: Out-of-sample model validation.

curves is similar, and the values themselves are very close in some cases. This is an indication of a good predictive model.

#### 4. Conclusion

We observed that Ghana's COVID-19 daily death count, from the very first day the pandemic was discovered in the country, is inflated with zeros (no deaths). This excessive number of zeros lead to overdispersion.

The trend of COVID-19 deaths per day in Ghana is characterized by a general increase from the onset of the pandemic in the country to about day 160 after which there is a general decrease onwards. The continuous decrease in the death toll amidst rise in daily infections and continuous disregard of safety protocols recently ought to be investigated.

We fitted a zero-inflated Poisson autoregressive model and zero-inflated negative binomial autoregressive model to the data in the partial-likelihood framework. The zero-inflated negative binomial autoregressive model outperformed the zero-inflated Poisson autoregressive model. We further obtained dynamic versions of the zero-inflated models. The dynamic zero-inflated Poisson autoregressive model, however, performed better than the dynamic negative binomial autoregressive model. Both dynamic models predicted an AR (1).

The predicted new deaths based on the ZINBA model showed that Ghana's COVID-19 deaths per day will rise sharply few days after 30<sup>th</sup> November 2020 and drastically fall just like that of the observed data.

#### Data Availability

The data used are made up of daily COVID-19 death count in Ghana from 13<sup>th</sup> March 2020 to 22<sup>nd</sup> November 2020

from Our World in Data, an official website for all COVID-19-related deaths (<https://ourworldindata.org/coronavirus-source-data>).

#### Conflicts of Interest

The authors declare no conflicts of interest.

#### Authors' Contributions

KT, IWA, and KAA sensed the idea. KT and IWA proposed the statistical methodology. IWA performed the statistical analysis. KT drafted the manuscript. KAA reviewed the manuscript. All authors agree to be answerable to all aspects of the work and jointly own the work. All authors read and approved the final manuscript.

#### References

- [1] Ghana Health Service, "Ghana confirms first two COVID-19 cases," 2020, [https://www.ghanahealthservice.org/covid19/downloads/covid\\_19\\_first\\_confirmed\\_GH.pdf](https://www.ghanahealthservice.org/covid19/downloads/covid_19_first_confirmed_GH.pdf).
- [2] Ghana Health Service, "COVID-19 updates," 2020, <https://www.ghanahealthservice.org/covid19/archive.php>.
- [3] H. A. Bonful, A. Addo-Lartey, J. M. K. Aheto, J. K. Ganle, B. Sarfo, and R. Aryeetey, "Limiting spread of COVID-19 in Ghana: compliance audit of selected transportation stations in the Greater Accra region of Ghana," *PLoS One*, vol. 15, no. 9, Article ID e0238971, 2020.
- [4] M. Sperrin and B. McMillan, "Prediction models for covid-19 outcomes," *BMJ*, vol. 371, 2020.
- [5] M. Maleki, M. R. Mahmoudi, M. H. Heydari, and K.-H. Pho, "Modeling and forecasting the spread and death rate of coronavirus (COVID-19) in the world using time series models," *Chaos, Solitons & Fractals*, vol. 140, Article ID 110151, 2020.
- [6] M. Daniyal, R. O. Ogundokun, K. Abid, M. D. Khan, and O. E. Ogundokun, "Predictive modeling of COVID-19 death cases in Pakistan," *Infectious Disease Modelling*, vol. 5, pp. 897–904, 2020.
- [7] IHME COVID-19 Forecasting Team, R. M. Barber, R. C. Reiner et al., "Modeling COVID-19 scenarios for the United States," *Nature Medicine*, vol. 27, 2020.
- [8] D. Dwomoh, S. Iddi, B. Adu et al., "Mathematical modeling of COVID-19 infection dynamics in Ghana: impact evaluation of integrated government and individual level interventions," *Infectious Disease Modelling*, vol. 6, pp. 381–397, 2021.
- [9] J. K. K. Asamoah, Z. Jin, B. Seidu, G. Q. Sun, F. T. Oduro, and F. Alzahrani, *A Mathematical Model and Sensitivity Assessment of COVID-19 Outbreak in Ghana and Egypt*, SSRN, New York, NY, USA, 2020.
- [10] M. Alqawba, N. Diawara, and N. Rao Chaganty, "Zero-inflated count time series models using Gaussian copula," *Sequential Analysis*, vol. 38, no. 3, pp. 342–357, 2019.
- [11] M. T. Hasan, G. Sneddon, and R. Ma, "Regression analysis of zero-inflated time-series counts: application to air pollution related emergency room visit data," *Journal of Applied Statistics*, vol. 39, no. 3, pp. 467–476, 2012.
- [12] K. Tawiah, S. Iddi, and A. Lotsi, "On zero-inflated hierarchical Poisson models with application to maternal mortality data," *International Journal of Mathematics and Mathematical Sciences*, vol. 2020, Article ID 1407320, 8 pages, 2020.



- [13] P. Wang, "Markov zero-inflated Poisson regression models for a time series of counts with excess zeros," *Journal of Applied Statistics*, vol. 28, no. 5, pp. 623–632, 2001.
- [14] M. Yang, *Statistical models for count time series with excess zeros*, Ph.D. thesis, University of Iowa, Iowa, IA, USA, 2012.
- [15] M. Yang, G. K. D. Zamba, and E. Joseph, J. E. Cavanaughb, "Markov regression models for count time series with excess zeros: a partial likelihood approach," *Statistical Methodology*, vol. 14, pp. 26–38, 2013.
- [16] D. Lambert, "Zero-inflated Poisson regression, with an application to defects in manufacturing," *Technometrics*, vol. 34, no. 1, pp. 1–13, 1992.
- [17] K. Lee, Y. Joo, J. J. Song, and D. W. Harper, "Analysis of zero-inflated clustered count data: a marginalized model," *Computational Statistics & Data Analysis*, vol. 55, no. 1, pp. 824–837, 2011.
- [18] N. M. Laird and J. H. Ware, "Random-effects models for longitudinal data," *Biometrics*, vol. 38, no. 4, pp. 963–974, 1982.
- [19] Y. Min and A. Agresti, "Random effect models for repeated measures of zero-inflated count data," *Statistical Modelling*, vol. 5, no. 1, pp. 1–19, 2005.
- [20] M. Ridout, C. G. B. Dem'etrio, and J. Hinde, "Models for count data with many zeros," in *Proceedings of the XIXth International Biometric Conference*, International Biometric Society, Cape Town, South Africa, July 1998.
- [21] K. K. W. Yau, K. Wang, and A. H. Lee, "Zero-inflated negative binomial mixed regression modeling of over-dispersed count data with extra zeros," *Biometrical Journal*, vol. 45, no. 4, pp. 437–452, 2003.
- [22] F. Zhu, "Zero-inflated Poisson and negative binomial integer-valued GARCH models," *Journal of Statistical Planning and Inference*, vol. 142, no. 4, pp. 826–839, 2012.
- [23] B. Kedem and K. Fokianos, *Regression Models for Time Series Analysis*, Wiley, Hoboken, NJ, USA, 2002.
- [24] H. Akaike, "A new look at the statistical model identification," *IEEE Transactions on Automatic Control*, vol. 19, no. 6, pp. 716–723, 1974.
- [25] K. P. Burnham and D. R. Anderson, "Multimodel inference," *Sociological Methods & Research*, vol. 33, no. 2, pp. 261–304, 2004.
- [26] K. Takeuchi, "Distribution of information statistics and criteria for adequacy of models," *Mathematical Science*, vol. 153, pp. 12–18, 1976.
- [27] R Core Team, *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria, 2019.
- [28] C. Kleiber and A. Zeileis, "Countreg: tools for count data regression," in *Proceedings of the R User Conference, useR!*, Brussels, Belgium, July 2017.
- [29] C. H. Weiss, *An Introduction to Discrete-Valued Time Series*, John Wiley & Sons Inc., Hoboken, NJ, USA, 2018.