Effects of climate variables on the COVID-19 mortality in Bangladesh

Rezaul Karim · Nazmin Akter

Received: 16 June 2022 / Accepted: 12 September 2022 / Published online: 18 October 2022
© The Author(s), under exclusive licence to Springer-Verlag GmbH Austria, part of Springer Nature 2022

Abstract
Infectious diseases such as severe acute respiratory syndrome (SARS) and influenza are influenced by weather conditions. Climate variables, for example, temperature and humidity, are two important factors in the severity of COVID-19’s impact on the human respiratory system. This study aims to examine the effects of these climate variables on COVID-19 mortality. The data are collected from March 08, 2020, to April 30, 2022. The parametric regression under GAM and semiparametric regression under GAMLSS frameworks are used to analyze the daily number of death due to COVID-19. Our findings revealed that temperature and relative humidity are commencing to daily deaths due to COVID-19. A positive association with COVID-19 daily death counts was observed for temperature range and a positive association for humidity. In addition, one-unit increase in daily temperature range was only associated with a 1.08% (95% CI: 1.06%, 1.10%), and humidity range was only associated with a 1.03% (95% CI: 1.02%, 1.03%) decrease in COVID-19 deaths. A flexible regression model within the framework of Generalized Additive Models for Location Scale and Shape is used to analyze the data by adjusting the time effect. We used two adaptable predictor models, such as (i) the Fractional polynomial model and (ii) the B-spline smoothing model, to estimate the systematic component of the GAMLSS model. According to both models, high humidity and temperature significantly (and drastically) lessened the severity of COVID-19 death. The findings on the epidemiological trends of the COVID-19 pandemic and weather changes may interest policymakers and health officials.

Keywords COVID-19 · Poisson regression · Negative binomial regression · GAMLSS · Semiparametric regression

1 Introduction
Coronavirus disease 2019 (COVID-19), caused by a novel coronavirus, has spread around the world and has become a severe public health issue (see Li et al. 2020). Almost 6.15 million deaths have been reported worldwide as of March 2022 (see Worldometers 2022). COVID-19 is rapidly spreading in many western and European temperate countries, such as Italy, France, Germany, Spain, the USA, and the UK, where temperatures range from 3 to 17 degrees Celsius, which is similar to Wuhan, China (see Bukhari and Jameel 2020). However, the growth rate of confirmed cases appears to be slower in Asian tropical countries such as Indonesia, Malaysia, Vietnam, Singapore, Thailand, and others (see Worldometers 2022), the majority of which are low- and middle-income countries (LMICs) with limited detection and response capabilities and have not implemented strict quarantine measures. As a result, on January 30, 2020, the World Health Organization (WHO) declared the SARS-CoV-2 outbreak a Public Health Emergency of International Concern, and on March 11, 2020, it was declared a global pandemic. Although the exact location of the outbreak is unknown, many early cases of COVID-19 have been linked to visitors to the Huanan Seafood Wholesale Market in Wuhan, Hubei, China (see Sun et al. 2020).

Bangladesh is well-known for its high population density and complex climate conditions, making it a climate-vulnerable country. Despite its large population (163 million) and the higher number of death than other countries, Bangladesh appears to be less severe. In Bangladesh, the first coronavirus cases were confirmed on March 08, 2020, by the country’s epidemiology institute, the Institute of Epidemiology Disease Control and Research (IEDCR); as of March 2020, approximately 29,119 death have occurred due to COVID-19.
in this country (see Bangladesh Covid 2020). It has been reported that the temperature, humidity, wind, and precipitation may favor the spread or the inhibition of epidemic episodes. Cheval et al. (2020) and Bashir et al. (2020) reported that the transmission of viruses is influenced by weather conditions and the density of people. As a result, there has been much discussion about whether rising temperatures at the start of spring and summer will result in reduced transmission in tropical countries like Bangladesh, as is the case with many viral respiratory infections (see, for example, Imai et al. 2014). There is also strong experimental and epidemiological evidence that the COVID-19 cases have tended to cluster in cooler, drier climates (for instance, Wang et al. 2020). However, there is still a lack of evidence, as some studies found no relationship between COVID-19 transmission and temperature (see, for example, Xie and Zhu 2020, Yao et al. 2020).

Furthermore, we know that viruses evolve over time, and SARS-CoV-2 is no exception. SARS-CoV-2 has been mutating at a rate of about 1–2 mutations per month (Callaway 2020). Depending on where in the SARS-CoV-2 virus the genome misconstructions occur, mutations can negatively or positively impact the virus’s ability to sustain and replicate. The researcher warned that these mutant SARS-CoV-2 genealogies would result in uncontrolled SARS-CoV-2 transmission in many parts of the world. Moreover, Bangladeshi virologists have discovered a new SARS-CoV-2 strain that is similar to the one discovered recently in the UK (The Daily Star 2020). We do not know how temperature and humidity affect the transmission of the SARS-CoV-2 strain after the mutation. As a result, it is critical to comprehend SARS-CoV-2 transmission behavior in light of current data.

This prompted us to launch a meteorological investigation into the impact of the weather on the COVID-19 outbreak. The main objective of this research is to investigate the effect of climate variables such as temperature and humidity on COVID-19 death by using parametric regression and semiparametric regression analysis. In all analyses, we consider the number of daily COVID-19 death as a response variable, and temperature and humidity are climate variables. Using daily death data from March 8, 2020, to January 31, 2021, it is demonstrated in the literature that temperature and humidity are highly significant in the daily number of COVID-19 deaths at a 5% level of significance (Karim et al. 2022). Based on more than two years of data, we are now attempting to decipher the seasonal patterns of COVID-19 transmission and COVID-19 spread from 8 March 2020 to 30 April 2022. This study highlights the importance of forward planning, preparation and decision-making to tackle this situation. With prompt, supportive, and empathetic collaboration among the government, citizens, and health experts, as well as international assistance, the country can minimize the impact of the pandemic. In the next section, we will go over the methods used to collect data and create statistical models in detail. A summary of the data analysis and findings are explained in Section 3.

2 Materials and methods

2.1 Data source

Data on COVID-19 cases are collected from the daily reports of the Institute of Epidemiology Disease Control and Research (IEDCR), Dhaka, Bangladesh, during the period of March 08, 2020, to April 30, 2022. Data are available on the website with the link https://en.wikipedia.org/wiki/COVID-19_pandemic_in_Bangladesh. The data is also available in https://covid19.who.int/region/searo/country/bd. The daily temperature (measured in °C) and humidity (%) of Bangladesh are collected from the website https://www.timeanddate.com/weather/bangladesh/dhaka.

To meet our goal, we have used the parametric regression models under the Generalized Additive Models (GAM) framework and the semiparametric regression models under the Generalized Additive Models for Location, Scale, and Shape (GAMLSS) framework. The best response distribution is chosen based on the minimum BIC under the GAMLSS modeling framework and We used two flexible predictor models to estimate the systematic part of the GAMLSS model, namely a fractional polynomial model and a B-spline smoothing model.

2.2 Generalized additive model

Generalized Linear Models (GLM) and Generalized Additive Models (GAM) respectively introduced by Nelder and Wedderburn (1972) and Hastie and Tibshirani (1990), are very popular in statistical data analysis. The GAM is an additive modeling technique where the impact of the predictive variables is captured through smooth functions, which can be nonlinear depending on the underlying patterns in the data. GAMs, in general, have the interpretability benefits of GLMs because the contribution of each independent variable to the prediction is clearly encoded. When any model includes nonlinear effects, GAM provides a regularized and interpretable solution, whereas other methods typically lack at least one of these three features. However, it has significantly more flexibility because the relationships between the independent and dependent variables are not assumed to be linear.

For response variable \( Y \), an exponential family of distributions such as Negative binomial and Poisson distributions is specified, along with a link function (such as log functions) that connects the expected value of \( Y \) to the predictor variables. The detailed procedure for fitting the GAM model is given in the following.
2.2.1 Poisson regression model

The foundation for Poisson regression is the Poisson distribution, which represents the distribution of the count response variable. The Poisson distribution is a discrete distribution that only has a probability value for non-negative integers; this property makes it an excellent choice for modeling count outcomes. The probability mass function (pmf) for the Poisson distribution,

\[ \Pr(Y = y | \mu) = \frac{\mu^y}{y!} e^{-\mu}; \quad y = 0, 1, 2, \ldots \]

where, \( \mu \) is the arithmetic mean of the number of occurrences that only takes discrete value. The parameter \( \mu \) defines the distribution’s mean and variance; both the mean and variance are equal \( \mu \), termed as equidispersion, that is \( \mu = E(Y) = \text{Var}(Y) \). Poisson regression is a type of regression analysis that uses a generalized linear model to model count data and contingency tables such as the arrival of a telephone call at a call center (see, for example, Dobson and Barnett 2018, Nelder and Wedderburn 1972). Poisson regression assumes that the response variable \( Y \) has a Poisson distribution and that a linear combination of unknown covariates can model the logarithm of its expected value. Therefore, the Poisson regression model is also known as a log-linear model (Loomis et al. 2005).

The mathematical model of Poisson regression is represented as

\[ \log(\mu) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_p X_p \]  

(1)

where \( \mu \) is the average count on the outcome variable given the specific values on the predictors \( X_1, X_2, \ldots, X_p \). It models the log of the expected count as a function of the predictor variables because the dependent variable is a count variable.

2.2.2 Negative binomial regression model

Negative Binomial regression can be used for overdispersed count data, that is, when the conditional variance exceeds the conditional mean. It is a generalization of Poisson regression that relaxes the Poisson model’s restrictive requirement that variance equals mean. The Poisson-gamma mixture distribution is the foundation of the standard negative binomial regression model. This approach is popular because it uses a gamma distribution to model Poisson heterogeneity. A gamma noise variable can be added to the Poisson distribution to make it more generic where the mean is 1 and a scale parameter is \( \nu \). The Negative Binomial (Poisson-gamma mixture) distribution that results is:

\[ \Pr(Y = y | \mu_i, \alpha) = \frac{\Gamma(y_i + \alpha^{-1})}{\Gamma(y_i + 1)(\Gamma(\alpha^{-1})} \left( \frac{\alpha^{-1}}{\alpha^{-1} + \mu_i} \right)^{\alpha^{-1}} \left( \frac{\mu_i}{\alpha^{-1} + \mu_i} \right)^{y_i} \]

where,

\[ \mu_i = t_i \mu \quad \text{and} \quad \alpha = \frac{1}{\nu}. \]

The parameter \( \mu \) is the mean incidence rate of \( y \) per unit of exposure (time, space, distance, area, or population size) which may be interpreted as the risk of a new occurrence of the event during a specified exposure period, \( t \) when no exposure given; it is assumed to be one. The negative binomial regression model can be expressed as

\[ \log \left( \frac{\mu_i}{t_i} \right) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_p X_p. \]  

(2)

The regression coefficients \( \beta_0, \beta_1, \beta_2, \ldots, \beta_p \) are unknown parameters of a set of \( p \) regressors that are estimated from a set of data.

2.2.3 Quasi-poison regression model

The Quasi-Poisson Regression is a generalization of the Poisson regression, which is used to model an overdispersed count variable (see Robert 1974). The Poisson model assumes that the variance equals the mean, which is not always the case. A Quasi-Poisson model, which assumes that the variance is a linear function of the mean, is more appropriate when the variance is greater than the mean. The Quasi-Poisson model is characterized by the first two moments (mean and variance) (Robert 1974) and the second parameter, often known as the overdispersion scaling parameter, \( \phi \), which is used in the estimation of the conditional variance. In addition, this model does not necessarily have a distributional form. However, it is demonstrated in the literature that how to generate a distribution for this model where reparameterization was required (Efron 1986). The model that was estimated with this correction now assumes a Poisson error distribution with mean and variance.

The model estimated with this correction now assumes a Poisson error distribution with mean \( \mu \) and variance \( \phi \mu \). If there is overdispersion in the data, the scaling parameter \( \phi \) will be greater than one; if there is equidispersion, the scaling parameter \( \phi \) will be equal to one, and the resulting model will be equivalent to the standard Poisson regression model. Finally, if the data are underdispersed, \( \phi \) will be less than one. The calculation of the scaling parameter is given by

\[ \phi = \frac{\chi^2_{\text{pearson}}}{\text{df}}. \]

Because the conditional variances are larger than their corresponding conditional means in the overdispersed model, the standard errors (based on the conditional
variances) are a factor of \( \sqrt{\phi} \) larger than the standard errors in the standard Poisson model.

### 2.3 Generalized additive models for location scale and shape

Rigby and Stasinopoulos (2005) proposed a generalized additive model for location, scale and shape (GAMLSS) to overcome some of the limitations associated with GLM and GAM models for regression analysis. It is a general framework of (semi)parametric regression models where the distribution of response variable does not necessarily belong to the exponential family and includes highly skewed and kurtotic continuous and discrete distribution. We consider the “daily number of death” due to COVID-19 as a response variable of the GAMLSS model. In the sequel, we denote, for notational convenience, “number of COVID-19 death” as “daily number of death”. For the response variable, we fit the GAMLSS model. The probability distribution of the response variable \( (Y) \) under the GAMLSS modeling framework is chosen based on the minimum Bayesian information criterion (BIC) and Akaike information criterion (AIC) values. The Beta Negative Binomial (BNB) distribution is selected for \( Y = \) daily number of death. A detailed selecting procedure is described in Section 3.2.

#### 2.3.1 Beta negative binomial distribution

The probability density function of the Beta Negative Binomial Distribution denoted by \( \text{BNB}(\mu, \sigma, \nu) \), is defined by

\[
f_{Y}(y|\mu, \sigma, \nu) = \frac{\Gamma(y + \frac{1}{\nu}) B(y + \frac{\mu}{\sigma}; \frac{1}{\nu} + \frac{1}{\nu} + 1)}{\Gamma(y + 1)(\Gamma(\frac{1}{\nu}) B(\frac{\mu}{\sigma}; \frac{1}{\nu} + 1))}
\]

for \( y = 0, 1, 2, \ldots \) and \( \mu > 0, \sigma > 0, \nu > 0 \) where, the location parameter \( \mu \) is the mean of \( Y \) and \( V \alpha(Y) = \mu \sigma(\nu + \sigma) \) \((1 + \mu \sigma)/(1 - \sigma)\) for detailed density can be found in Wang (2011). We are interested in estimating the mean function in the regression settings.

The covariates for both response variables are time (in days), temperature, and humidity are considered for this section. The beauty of the GAMLSS model is that the systematic part of it can be elaborated to support modeling not only the location (typically, mean), but also other distribution parameters such as scale and shape. These parameters could be linear parametric functions of covariates and/or random effects, or they could be additive non-parametric functions of covariates and/or random effects. In this study, we will use flexible predictor models to find the smoothing function of the predictor time using fractional polynomial and B-spline functions. To estimate the conditional mean of the response variable \( Y \) as daily number of death and given covariate \( X = (\text{time}, \text{temperature}, \text{humidity}) \), we have to estimate the parameters (as a function of \( X \)) of the conditional distribution of \( Y \) given \( X \). Therefore, the flexible regression models for the location function \( \mu(X) \) and the scale function \( \sigma(X) \) under the flexible GAMLSS modeling framework can be written as

\[
\mu(X; \beta) = \beta_0 + f(\text{time}; \beta_1) + \beta_2 \times \text{temperature} + \beta_3 \times \text{humidity},
\]

and

\[
\log(\sigma(X; y)) = \gamma_0 + f(\text{time}; \gamma_1) + \gamma_2 \times \text{temperature} + \gamma_3 \times \text{humidity}.
\]

The (penalized) maximum likelihood estimation is used to estimate the parameters of the model (3) and (4).

#### 2.3.2 Flexible regression with fractional polynomial function

The fractional polynomial in flexible predictor models is a generalization of the polynomial function. The general form of a fractional polynomial in \( x \) of degree \( m \) can be written as

\[
f_p(x; \theta_1, \theta_2, \ldots , \theta_m) = \sum_{i=0}^{m} \theta_i H_i(x)
\]

where \( m \) is an integer and

\[
H_i(x) = \begin{cases} 
  x^{pi} & \text{if } pl \neq p_{i-1} \\
  H_{i-1}(x) \times \log(x) & \text{if } pl = p_{i-1},
\end{cases}
\]

with \( p_0 = 0 \) and \( H_0(x) = 1 \), for a sequence of powers \( p_1 \leq \ldots \leq \max(3, m) \).

The optimal combination of powers will be selected by using the smallest value of BIC. Now, for the response variable of daily number of death, we select \( p_1 = -1, p_2 = 3 \) and \( p_3 = 3 \) with \( m = 3 \) and the fractional polynomial in (days) variable for the model (5) can be written as (5) can be written as

\[
f_p(\text{time}; \theta_1, -1, 3, 3) = \theta_{10} + \theta_{11} (\text{time})^{-1} + \theta_{12} (\text{time})^3 + \theta_{13} (\text{time})^3 \log(\text{time}).
\]

#### 2.3.3 Flexible smoothing regression with B-splines model

Flexible smoothing function with basis spline (B-spline) is also fitted in order to get a more flexible approximation to the data. A general form of B-spline predictor model of \( x \) for the degree \( D \) can be written as

\[
f_b(x; \theta_0, D, K) = \sum_{j=0}^{D} \theta_0 y^j + \sum_{k=D+1}^{D+K} \theta_k (x - b_k)^D H(x > b_k),
\]

where \( K \) is the degree of freedom.
where $K$ is the number of knot values, $b_k$ is the knot value at $k$th interval or piece and $H(x > b_k)$ is the Heaviside function taking value 1 if $x > b_k$, otherwise 0. The combination of $D$, $K$, and the number of knots values will be chosen based on the lowest value of BIC.

## 3 Data analysis and results

### 3.1 Exploratory data analysis

Due to COVID-19 in Bangladesh, the total number of death is 27514 from 8 March 2020 to 30 April 2022 which shows a positive correlation with daily highest Temperature (Pearson’s $r = 0.228$) and Humidity (Pearson’s $r = 0.295$).

The descriptive statistics of number of COVID-19 death and climatology parameters for 764 days are summarized in Table 1. The average number of daily new death due to COVID-19 is almost 36 whereas the average temperature and humidity are 30.30°C and 63.67%, respectively. During this experimental period, the highest temperature was 37°C and the lowest temperature was 10°C whereas the lowest humidity was 21% and the highest humidity was 100%.

Figure 1 shows a histogram with a kernel density plot of the daily number of deaths caused by COVID-19. It demonstrates that the distributional shape of the number of deaths caused by COVID-19 appears symmetric, implying that the bell-shape distribution is one of the best probability models for this variable, whereas it reveals that the distributional shape of the number of death cases due to COVID-19 looks similar to a skewed pattern, indicating a skewed distribution would be more suitable for predicting this variable’s values.

The scatter plots of the number of death due to COVID-19 against daily temperature, humidity and time for the period from March 08, 2020, to April 30, 2022, are drawn respectively in Fig. 2. These graphs show that there is a nonlinear relationship between the response variables and covariates.

Without adjusting the time variable in the model, we consider the following regression model to explore only the conditional relationship between $Y = \text{daily number of death}$ and two covariates, namely temperature and humidity. The Classical regression model is, for $i = 1, 2, ..., n$

$$y_i = \beta_0 + \beta_1 \times \text{temperature}_i + \beta_2 \times \text{humidity}_i + \epsilon_i \quad (8)$$

where the response variables ($y_i$) is the number of death and $\epsilon_i$ is the disturbance term for $i$th individual. Under the classical regression model assumptions (see, for example, Gujarati and Porter 2010), the summary statistics of the model (8) are tabulated in Table 2. The exploratory results show that the temperature and humidity are highly significant on $Y$. Table 2 shows that both coefficients are positively associated.

We next consider the time (in days) variable as a covariate in the model. Since the exploratory data analysis shows a nonlinear relationship between time and response variables,
The Generalized Poisson Regression Model (Negative Binomial regression model, Quasi-Poisson regression model) can be employed to handle overdispersion. These Generalized Poisson regression model also is designed to analyze count data. Quasi-Poisson and Negative Binomial regression models have equal numbers of parameters (two parameters), though the variance of a Quasi-Poisson model is a linear function of the meanwhile the variance of a negative binomial model is a quadratic function of the mean (see, for example, Hoef and Boveng 2007). However, Negative binomial regression models do not assume equal mean and variance and account for overdispersion in the data, which occurs when variance exceeds the conditional mean (see Piza 2012). The Negative Binomial and Quasi-Poisson models are the most commonly used because they are widely available in software, and they generalize easily to the regression case, which we outline in the subsequent Sections 2.2.2, and 2.2.3 respectively.

### 3.1.1 Poisson regression analysis under GAM framework

The Poisson regression coefficient can be interpreted as follows: The difference in the logs of expected counts is expected to change by the respective regression coefficient for a one-unit change in the predictor variable, assuming the other predictor variables in the model remain constant. In other words, the Poisson regression coefficient \( \beta \) associated with a predictor \( X \) is the expected change, on the log scale, in the outcome \( Y \) per unit change in \( X \)'s. So holding all other variables in the model constant, increasing \( X \) by 1 unit multiplies the rate of \( Y \) by \( e^\beta \). The Poisson regression model for the response variable \( Y = \) daily number of death due to COVID-19 given in (1) is estimated within the GLM modeling framework. The fitted predictor model for the \( \hat{\mu} = E(\text{daily number of death}) \) is:

\[
\log(\mu) = \beta_0 + \beta_1 \times \text{temperature} + \beta_2 \times \text{humidity}. \tag{9}
\]

and confidence interval can be calculated using the following formula:

\[
95\% \text{ Confidence Interval} = \exp(\hat{\beta} \pm 2 \times SE)
\]

The summary statistics of the model (9) is tabulated in Table 3.

From Table 3 we found that the coefficient of the \( \beta \)'s are statistically significant (associated with a \( p \)-value < 0.001), and the dispersed parameter fixed at \( \phi = 1 \). Therefore, the temperature and humidity impact the rate of COVID-19’s death. The sign of the estimated value of both parameters is positive. It indicates that high temperature and humidity may increase the death rate. The equation of the Poisson model (9) can be rewritten as:

\[
\log(\mu) = -1.948 + 0.119 \times \text{temperature} + 0.029 \times \text{humidity}. \tag{10}
\]

Then \( e^{\beta_1} = e^{0.119} = 1.126 \) is the Rate Ratio which means temperature is associated with an increase of approximately 12\% (1.126 − 1 = .126) in the death rate. Similarly, \( e^{\beta_2} = e^{0.029} = 1.029 \) is the Rate Ratio which means humidity is positively associated with an increase of almost 3.0\% (1.029 − 1 = 0.029) in the death rate due to COVID-19.

### 3.1.2 Negative binomial regression analysis under GAM framework

In the estimated Negative Binomial regression coefficients for the model (2) the dependent variable \( Y \) is a count variable that is either overdispersed or underdispersed, and the model models the log of the expected count as a function of the predictor variables. We can interpret the Negative Binomial regression coefficient as follows: for a one-unit change in the predictor variable, the difference in the logs of expected counts of the response variable is expected to change by the respective regression coefficient, given the other predictor variables in the model are held constant.

The summary output of this model is shown in Table 4.

This Table 4 provides the estimated regression model and associated statistics. The temperature has a coefficient of 0.081, which is statistically significant. It means for each one-unit increase per degree in temperature, the expected log count of the daily number of death increases by 0.081. Similarly, the humidity is also significant having a coefficient 0.025 which means for each one-unit increase in humidity, the expected log count of the daily number of death

### Table 2 Summary statistics of the estimated OLS model

| Estimate       | \( \hat{\beta} (se(\hat{\beta})) \) | \( t \)-value | \( P \)-value |
|----------------|-------------------------------------|--------------|--------------|
| Intercept      | \(-120.165 (15.163)\)               | \(-7.925\)   | <0.001       |
| Temperature    | \(3.172 (0.437)\)                   | 7.251        | <0.001       |
| Humidity       | \(0.942 (0.102)\)                   | 9.169        | <0.001       |

### Table 3 Summary statistics of the estimated Poisson regression model

| Estimate       | \( \hat{\beta} (se(\hat{\beta})) \) | \( z \)-value | \( P \)-value | 95% CI      |
|----------------|-------------------------------------|--------------|--------------|-------------|
| Intercept      | \(-1.948 (0.073)\)                 | \(-26.42\)   | <0.001       | (0.123, 0.164) |
| Temperature    | \(0.119 (0.001)\)                  | 57.26        | <0.001       | (1.117,1.121) |
| Humidity       | \(0.029 (0.001)\)                  | 71.75        | <0.001       | (1.027,1.031) |
increases by 0.025%. According to the estimated value the model (2) can be rewritten as:

\[ \hat{\mu} = \exp[-0.647 + 0.081 \times \text{temperature} + 0.025 \times \text{humidity}] \quad (11) \]

the mean value is 36.013 and the variance is 2498.82. Hence the dispersion parameter \( \theta \) is 0.845. A large-sample confidence interval for the values of the coefficient can be determined as

95% Confidence Interval = \( \exp(\beta \pm 2 \times SE) \) \quad (12)

A relative risk of one implies there is no difference in the event if the exposure has or has not occurred. Relative risk provides an increase or decrease in the likelihood of an event based on some exposure. It has the benefit of being a ratio of risks which means it can be applied to populations with differing disease prevalence. From Table 5 we got the relative risk is greater than 1 for both exposure (temperature and humidity), then the death due to COVID-19 may more likely to occur.

### 3.1.3 Quasi-poison regression analysis under GAM framework

The coefficient interpretation for the overdispersed Poisson model is the same as for the standard Poisson model. The scaling factor also affects the deviance for this model; the deviance for the overdispersed Poisson model is equal to the deviance for the standard Poisson model divided by \( \phi \). This model’s lower deviance indicates a better fit (for instance, Coxe et al. 2009).

Table 6 shows the regression coefficients for the Quasi-Poisson model. The square root of the overdispersion parameter is \( \sqrt{\phi} = 6.395 \). The value of \( \phi \) is 35.825, which is substantially larger than 1.00 for the standard Poisson model that assumes equidispersion: The data clearly exhibit overdispersion the intercept represents the log mean daily number of death when all predictor variables in the model are equal to zero. For instance, temperature = 0. If we exponentiate the intercept, then \( \exp(\beta_0) \) represents the mean of the daily number of death when all predictor variables are equal to zero. In practice, the intercept is unknown and estimated from the data. So the estimated value of \( \beta_0 \) is −1.948. The coefficient of the \( \beta \)'s is statistically significant (associated with a \( p \)-value < 0.001). Therefore temperature and humidity show a positive impact on the rate of COVID-19’s death because its sign is positive. Additionally, the comparison between Negative Binomial model and the Quasi-Poisson model is illustrated through the model diagnosis diagram of Fig. 5 in the Appendix section.

### 3.2 Data exploration for generalized additive models for location scale and shape (GAMLSS)

For selecting the best probability model for the response variable \( Y = \) daily number of death, the summary including AIC and BIC values with their degrees of freedom of all selected candidate distributions coming from the GAMLSS family, are provided in Table 9 in the Appendix. Above all of the distributions, we selected five possible candidate distributions based on the minimum BIC provided in Table 10 in the Appendix. It is noticed that the smallest BIC and AIC are observed for the Beta Negative Binomial (BNB) (discrete) model. In contrast, the highest value of BIC (and also AIC) is observed for the Poisson Inverse Gaussian model. Based on the minimum BIC, we select the BNB model to explain the severity of SARS-CoV-2 death for further investigation.

### Table 4 Summary statistics of the estimated Negative Binomial regression model

| Estimate     | \( \hat{\beta} \) | se(\( \hat{\beta} \)) | z-value | P-value |
|--------------|-------------------|------------------------|---------|---------|
| Intercept    | −0.647            | 0.365                  | −1.777  | 0.075   |
| Temperature  | 0.081             | 0.011                  | 7.664   | <0.001  |
| Humidity     | 0.025             | 0.002                  | 10.525  | <0.001  |

### Table 5 Relative Risk for each independent variable in Negative Binomial regression

| Estimate     | Coefficients (\( \beta \)) | Relative Risk(\( \text{RR} \)) | Lower 95% Confidence interval | Upper 95% Confidence interval |
|--------------|-----------------------------|---------------------------------|-------------------------------|------------------------------|
| Intercept    | −0.647                      | 0.523                           | 0.252                         | 1.086                        |
| Temperature  | 0.081                       | 1.084                           | 1.060                         | 1.109                        |
| Humidity     | 0.025                       | 1.025                           | 1.021                         | 1.029                        |
3.2.1 Flexible regression with fractional polynomial function

A fractional polynomials flexible function for the “time” variable given in (6) is estimated within the GAMLSS modeling framework via the best-chosen probability distribution of the response variable. The fitted flexible predictor model for the $\mu(X)$ for the daily number of death is

$$\hat{\mu}(X; \hat{\beta}) = -0.923 + f_p(\text{time}, \hat{\beta}_i) + 0.078 \times \text{temperature} + 0.021 \times \text{humidity},$$

and the estimated flexible predictor model (4) is $\sigma(X; \hat{\gamma}) = \exp(-3.299) = 0.327$. We, here, leave out the insignificant term of the estimated model. The corresponding estimated fractional polynomial model for the $\mu(X)$ in time (in days) of degree 3 is

$$f_p(\text{time}, \hat{\beta}_i) = 0.894 - 0.463 \times (\text{time})^{-1} + 0.040 \times (\text{time})^3 - 0.236 \times (\text{time})^3 \times \log(\text{time}).$$

The summary statistics of this estimated flexible predictor model (13) is tabulated in Table 7. Hence, the estimated flexible regression model for the mean function $E(Y|X) = \mu(X)$ of the conditional BNB distribution under the GAMLSS modeling framework is

$$\hat{\mu}(X; \hat{\beta}) = -0.029 - 0.463 \times (\text{time})^{-1} + 0.040 \times (\text{time})^2 - 0.236 \times (\text{time})^3 \times \log(\text{time})$$

$$+ 0.078 \times \text{temperature} + 0.021 \times \text{humidity},$$

(14)

So we obtain the estimated flexible regression model for mean function $E(Y|X) = \mu(X)$ of conditional BNB under the GAMLSS modeling framework

$$E(Y|X) = \hat{\mu}(X; \hat{\beta})$$

$$= -0.029 - 0.463 \times (\text{time})^{-1} + 0.040 \times (\text{time})^2 - 0.236 \times (\text{time})^3 \times \log(\text{time})$$

$$+ 0.078 \times \text{temperature} + 0.021 \times \text{humidity},$$

note that the value of fixed parameters $\nu$ is 0.036($>0$) in the GAMLSS modeling framework. We found the Global Deviance is 6234.513, AIC is 6256.513, and BIC is 6307.537 for the final fitted model. Table 7 shows that the temperature and humidity are highly significant on the number of SARS-CoV-2 COVID-19 mortality. In addition, the regression coefficients for both temperature and humidity are positive which indicates that there is a positive relationship between these variables and the number of SARS-CoV-2 death cases.

### Table 7 Summary statistics of the estimated flexible predictor model

| Estimate | $\hat{\beta}$ | se(\$\hat{\beta}\$) | t-value | P-value |
|----------|---------------|-----------------|----------|----------|
| Intercept | $-0.923$      | 0.272           | $-3.385$ | $< 0.0001$ |
| Temperature | $0.078$     | 0.007           | $10.513$ | $< 0.0001$ |
| Humidity  | $0.021$       | 0.002           | $11.830$ | $< 0.0001$ |

3.2.2 Flexible smoothing regression with B-splines function

For the response variable daily number of death, we also use B-spline function given in (7) for estimating $\mu(X; \beta)$ and $\sigma(X; \gamma)$ of the BNB distribution. With $D = 3$ and $K = 5$ in the model (7), the B-spline predictor function for estimating $\mu(X; \beta)$, the estimated B-spline smoothing function of $f_p(\text{time}; \beta_0, 3, 5)$, for $i = 1, 2, ..., n$ is

@rcl@ $@f_p(\text{time}; \beta_0, 3, 5) = -0.235 + 6.322 \times \text{time} + 6.827 \times \text{time}^2$

$+ 5.827 \times \text{time}^3 + H(\text{time} > b_3)(3.950 \times (\text{time} - 128.5)^3$

$+ 11.122 \times (\text{time} - 256.0)^3 + 0.223 \times (\text{time} - 383.5)^3$

$+ 8.668 \times (\text{time} - 511.0)^3 - 2.156 \times (\text{time} - 638.5)^3$. 

(15)

With $D = 3$ and $K = 2$ in the model (7), the estimated function of $f_p(\text{time}; \beta_0, 3, 2)$ for $i = 1, 2, ..., n$ is

@rcl@ $@f_p(\text{time}; \beta_0, 3, 2) = 6.915 - 60.472 \times \text{time} + 2.441 \times \text{time}^2 - 8.855 \times \text{time}^3$

$- H(\text{time} > b_2)(4.416(\text{time} - 256.0)^3 + 8.824(\text{time} - 511.0)^3)$. 

(16)

Using the estimated B-spline function for estimating $\mu(X; \beta)$ given in (7), we find the estimated flexible regression function of $E(Y|X) = \mu(X; \beta)$ which is

@rcl@ $@\hat{\mu}(X; \hat{\beta}) = 0.235 + 6.322 \times \text{time} + 6.827 \times \text{time}^2$

$+ 5.827 \times \text{time}^3 + H(\text{time} > b_3)\{3.950 \times (\text{time} - 128.5)^3$

$+ 11.122 \times (\text{time} - 256.0)^3 + 0.223 \times (\text{time} - 383.5)^3$

$+ 8.668 \times (\text{time} - 511.0)^3 - 2.156 \times (\text{time} - 638.5)^3$

$- 0.061 \times \text{temperature} - 0.008 \times \text{humidity}.\}$

By using the estimated B-spline model, the estimated scale function $\hat{\sigma}(X; \hat{\gamma})$ for $i = 1, 2, ..., n$ is

@rcl@ $@\hat{\sigma}(X; \hat{\gamma}) = \exp(6.915 - 60.472 \times \text{time} + 2.441 \times \text{time}^2 - 8.855 \times \text{time}^3$

$- H(\text{time} > b_2)(4.416(\text{time} - 256.0)^3 + 8.824(\text{time} - 511.0)^3)$

$- 0.118 \times \text{temperature} + 0.006 \times \text{humidity}).\}$

The summary statistics of the estimated function $\hat{\mu}(X; \hat{\beta})$ and log ($\hat{\sigma}(X; \hat{\gamma})$) are presented in Table 8. The estimated $\nu$ coefficient is $\exp(-3.366)$ or 0.035 here. For this estimated model, the Global Deviance, AIC and BIC are 5347.333, 5387.333 and 5480.104, respectively. We also see the slope coefficient of temperature ($\beta_1$) and humidity ($\beta_2$) are negative, indicating a negative impact on the location function $\hat{\mu}(X; \hat{\beta})$. In addition, the regression coefficient for both temperature and humidity are highly significant on $\hat{\mu}(X; \hat{\beta})$. Similarly, for estimated log ($\hat{\sigma}(X; \hat{\gamma})$), we can see the slope coefficient of temperature ($\gamma_1$) is negative which indicates that there is a negative relationship between these variables and humidity ($\gamma_2$) coefficient is positive which indicates...
Effects of climate variables on the COVID-19 mortality in Bangladesh

We also calculate the predicted values of the response variable via fractional polynomial and B-spline models. The graphical presentation of actual values and predicted values is depicted in Fig. 3. From Fig. 3, we have seen that the estimated curve via the B-spline function is a smooth curve which is expected. On the other hand, the estimated curve via fractional polynomial function is not smooth. However, estimated both curves are very close.

4 Discussion

Since December 2019, a new epidemic of the SARS Coronavirus, known as SARS-CoV-2 or COVID-19, has emerged, causing millions of deaths around the world. The drastically spread of COVID-19, can be affected by a number of factors, including population density, migratory flow, host immunity, medical care quality and, presumably, climate conditions such as temperature and humidity (see Dalziel et al. 2018). It is thought that high temperatures and humidity have a combined effect on coronavirus inactivation. This combination has the potential to impair transmission, resulting in deaths. Since December 2019, a new epidemic of the SARS Coronavirus, known as SARS-CoV-2 or COVID-19, has emerged, causing millions of deaths around the world. The drastically spread of COVID-19, can be affected by a number of factors, including population density, migratory flow, host immunity, medical care quality and, presumably, climate variables (see Dalziel et al. 2018). It is thought that high temperatures and humidity have a combined effect on coronavirus inactivation. This combination has the potential to impair transmission, resulting in deaths. In this research we relied on the daily count of the number of COVID-19 death per day from Institute of Epidemiology Disease Control and Research (IEDCR), Dhaka, Bangladesh. A generalized additive model location scale and shape model is used to examine the effect of temperature and humidity on the number of confirmed SARS-CoV-2 daily death cases due to COVID-19 separately.

One of the purposes of this study was to analyze count outcomes with and without considering time variables by using appropriate statistical tools. We began by discussing to analyze count outcomes and used the Poisson regression under GLM modeling framework. We discussed the limitations of Poisson regression related to the assumption of equidispersion. Two models that relax the equidispersion constraint, the Quasi-Poisson model and the Negative Binomial model, were illustrated in this research. We expanded the COVID-19 death example to include these models. The estimated results and interpretations conclude that Negative Binomial regression model can overcome overdispersion.

Table 8 The summary statistics of flexible regression models of $\mu(X, \beta)$ and $\log(\sigma(X, \gamma))$ via B-spline smoothing function for the response variable number of death

| $\beta$ | $\tilde{\beta}(\tilde{\sigma}(\tilde{\gamma}))$ | $t$-value | $P$-value | $\gamma$ | $\tilde{\gamma}(\tilde{\sigma}(\tilde{\gamma}))$ | $t$-value | $P$-value |
|--------|-------------------------------------|------------|-----------|--------|-------------------------------------|------------|-----------|
| $\beta_{00}$ | $-0.235 (0.445)$ | $-0.529$ | 0.597 | $\gamma_{00}$ | $6.915 (2.439)$ | 2.835 | 0.004 |
| $\beta_{01}$ | $6.322 (0.402)$ | $15.705$ | <0.001 | $\gamma_{01}$ | $-60.472 (3.602)$ | $-16.787$ | <0.001 |
| $\beta_{02}$ | $6.827 (0.234)$ | $29.062$ | <0.001 | $\gamma_{02}$ | $2.441 (2.264)$ | 1.078 | 0.281 |
| $\beta_{03}$ | $5.827 (0.322)$ | $18.087$ | <0.001 | $\gamma_{03}$ | $-8.855 (3.097)$ | $-2.859$ | 0.004 |
| $\beta_{04}$ | $3.949 (0.299)$ | $13.183$ | <0.001 | $\gamma_{04}$ | $-4.416 (2.380)$ | $-1.855$ | 0.063 |
| $\beta_{05}$ | $11.122 (0.317)$ | $35.015$ | <0.001 | $\gamma_{05}$ | $-8.824 (2.781)$ | $-3.172$ | 0.001 |
| $\beta_{06}$ | $0.223 (0.342)$ | $0.652$ | <0.001 | | | | |
| $\beta_{07}$ | $8.668 (0.407)$ | $21.282$ | <0.001 | | | | |
| $\beta_{08}$ | $-2.156 (0.618)$ | $-3.487$ | <0.001 | | | | |
| $\beta_{1}$ | $-0.061 (0.007)$ | $-8.196$ | <0.001 | $\gamma_{1}$ | $-0.118 (0.028)$ | $-4.211$ | <0.001 |
| $\beta_{2}$ | $-0.008 (0.001)$ | $-5.949$ | <0.001 | $\gamma_{2}$ | $0.006 (0.008)$ | 0.747 | 0.455 |

Fig. 3 Fractional polynomial curve vs Basis spline curve of daily Number of death due to the COVID-19 vs. days during the period March 08, 2020, to April 30, 2022

a positive relationship between these variables also there both regression coefficients are significant on the number of COVID-19 death cases at 1% and 5% level of significance respectively.

We also calculated the predicted values of the response variable via fractional polynomial and B-spline models. The graphical presentation of actual values and predicted values is depicted in Fig. 3. From Fig. 3, we have seen that the estimated curve via the B-spline function is a smooth curve which is expected. On the other hand, the estimated curve via fractional polynomial function is not smooth. However, estimated both curves are very close.
problem with the appearance of the dispersion parameter. The interpretation of the results and tests of competing estimated death model is pretty straightforward. In addition we further analyzed the COVID-19 death data with considering time variable under GAMLSS framework. We used the flexible smoothing regression model to investigate the significant effects of temperature and humidity after adjusting the time variable. The best response distribution is chosen based on the minimum Bayesian information criterion (BIC) under the GAMLSS modeling framework. To estimate the systematic part of the GAMLSS model, we have employed two flexible predictor models such as (i) fractional polynomial model and (ii) B-spline smoothing model. Both models suggested that high temperature and high humidity may reduce the severity of COVID-19 death. Many researches support these results (see, for example, in Mecenas et al. 2020) but these are opposite of the findings of Gupta (2020). Karim et al. investigated the effects of temperature and humidity on the transmission of SARS-CoV-2 and the Spread Covid-19 using the daily number of SARS-CoV-2 infected new cases, and the number of death due to Covid-19 are considered the response variables (Karim et al. 2022). Using flexible regression model under the GAMLSS framework it was found that the temperature and humidity have a significant impact on the transmission of COVID-19. Particularly, the temperature is highly significant in the number of SARS-CoV-2 infected new cases and number of death due to COVID-19. In contrast, the humidity is significant on the number of SARS-CoV-2 infected new cases, but it is insignificant on the number of death due to COVID-19 at a 5% level of significance. This research also revealed that both the temperature and humidity inversely affected the daily number of deaths and new cases of COVID-19.

Moreover, in Term plot in Fig. 4, these changes are more obvious where, with the increases of temperature and humidity, the partial changes of death severity are slightly decreasing on the one hand, and on the other hand, with time, the death severity shows a seasonal variation.

5 Conclusions

To sum up, we have provided an introduction to regression models that are useful in analyzing count data. These models are attractive alternatives to analyses using the classical regression tool, particularly when the mean of the outcome variable is not large (less than 10 as a rule of thumb). The majority of recent studies in literature have relied on correlation analysis between climate indicators and epidemic data. Direct studies with COVID-19 in various environmental conditions, however, are still lacking. Our findings on the effect of temperature and humidity on the seasonal viability and severity of COVID-19 death were remarkably consistent. The findings of our research, which included several statistical concepts and models, suggest that climatic variables such as temperature and humidity may influence COVID-19 death. Warmer, humid climates appear to have a lower risk of COVID-19 death severity. Thus the gradually decreasing temperature may make epidemic control more difficult. However, these variables alone could not account for the majority of the variation in disease transmission. The countries most affected by the disease should prioritize health policies, even if their climates are less conducive to the virus. Public isolation policies, herd immunity, migration patterns, population density, and cultural factors could all play a role in the disease’s spread. As a result, weather conditions linked to health policies are extremely valuable information for the benefit of humanity at this critical time. Environmental research, in addition to epidemiology, biological sciences, and medical sciences, is urgently needed to control the spread of COVID-19 globally.

Fig. 4 Term plot. (a) Daily Number of death due to the COVID-19 vs. time. (b) Daily Number of death due to the COVID-19 vs. Daily temperature. (c) Daily Number of death due to the COVID-19 vs. daily humidity during the period March 08, 2020–April 30, 2022
Table 9 Distribution of the response variable “number of death” under the GAMLSS modelling framework

| Distribution                  | AIC     | BIC     | df  |
|-------------------------------|---------|---------|-----|
| Beta Negative Binomial        | 6256.513| 6307.537| 11  |
| Delaport                      | 6425.108| 6476.132| 11  |
| Discrete Burr II              | 6394.089| 6445.113| 11  |
| Double Poisson                | 6412.795| 6459.181| 10  |
| Exponential Gaussian          | 6911.149| 6952.896| 9   |
| Geometric                     | 6454.686| 6496.433| 9   |
| Geometric original            | 6454.686| 6496.433| 9   |
| Negative Binomial I           | 6421.215| 6467.600| 10  |
| Negative Binomial II          | 6373.351| 6419.737| 10  |
| Negative Binomial Family      | 6335.344| 6386.368| 11  |
| Normal Linear Quadratic       | 6772.800| 6819.186| 10  |
| Poisson                       | 18254.61| 18296.35| 9   |
| Poisson Inverse Gaussian      | 6401.376| 6447.761| 10  |
| Shash                         | 6714.273| 6769.936| 12  |
| Shash Original                | 6739.845| 6795.507| 12  |
| Yule                          | 9025.550| 9048.743| 5   |
| Zero Adjusted Poisson         | 18218.90| 18265.28| 10  |
| Zero Inflated Beta Negative Binomial | 6300.366| 6346.752| 10  |
| Zero Inflated Negative Binomial| 6423.217| 6474.241| 11  |
| Zero Inflated Poisson         | 18211.41| 18257.79| 10  |

Note: Five Distributions based on minimum BIC values are emphasized among all distributions

Table 10 Goodness-of-fit statistics for selecting the best response distribution of number of death cases

| Distribution                | AIC     | BIC     | m  | df |
|-----------------------------|---------|---------|----|----|
| Beta Negative Binomial      | 6256.513| 6307.537| 3  | 11 |
| Zero Inflated Beta Negative Binomial | 6300.366| 6346.752| 2  | 10 |
| Negative Binomial Family    | 6389.871| 6431.619| 2  | 9  |
| Negative Binomial II        | 6420.996| 6458.105| 2  | 8  |
| Discrete Burr II            | 6436.302| 6478.049| 2  | 9  |

Note: The distribution having the minimum BIC value is selected for the best response distribution
Acknowledgements The authors thank an Editor and two anonymous reviewers for their valuable suggestions and comments, which led to a considerable improvement of the manuscript.

Author contribution Md. Rezaul Karim designed and supervised the research. Nazmin Akter collected data and carried out the implementation. She performed the data analysis and wrote a draft copy of the manuscript. Md. Rezaul Karim checked the data analysis and results and then finalized the manuscript.

Data availability The data are available on the website, and the link is provided in Section 2. It will be provided if anyone requires this.

Code availability The R code is available. It will be provided if anyone requires this.

Declarations

Ethics approval We have conducted ourselves with integrity, fidelity, and honesty. We have not intentionally engaged in or participated in any form of malicious harm to another person or animal.

Consent for publication Not applicable
Consent to participate  Not applicable

Conflict of interest  The authors declare no competing interests.

References

Bangladesh Covid. update (2020) Institute of epidemiology, disease control and research (iedcr), 19
Bashir MF, Ma B, Komal B, Bashir MA, Tan D, Bashir M, et al. (2020) Correlation between climate indicators and covid-19 pandemic in new york, usa. Sci Total Environ 728:138835
Bukhari Q, Jameel Y (2020) Will coronavirus pandemic diminish by summer? Available at SSRN 3556998
Callaway E (2020) The coronavirus is mutating—does it matter?. Nature 585(7824):174–178
Cheval S, Adamescu CM, Georgiadis T, Herrnegger M, Piticar A, Legates DR (2020) Observed and potential impacts of the covid-19 pandemic on the environment. Int J Environ Res Public Health 17(11):4140
Coxe S, West SG, Aiken LS (2009) The analysis of count data: a gentle introduction to poisson regression and its alternatives. J Personality Assessment 91(2):121–136
Dalziel BD, Kissler S, Gog JR, Viboud C, Bjørnstad ON, Metcalf CJ, Grenfell Bryan T (2018) Urbanization and humidity shape the intensity of influenza epidemics in us cities. Science 362 (6410):75–79
Dobson AJ, Barnett AG (2018) An introduction to generalized linear models. Chapman and hall/CRC, Boca Raton
Efron B (1986) Double exponential families and their use in generalized linear regression. J Am Stat Assoc 81(395):709–721
Gujarati DN, Porter DC (2010) Essentials of econometrics
Gupta D (2020) Effect of ambient temperature on COVID-19 infection rate. Available at SSRN 3558470
Hastie TJ, Tibshirani RJ (1990) Generalized Additive Models, volume 43. CRC press, Boca Raton
Hoef JMV, Boveng PL (2007) Quasi-poisson vs. negative binomial regression: how should we model overdispersed count data?. Ecol 88(11):2766–2772
Imai WC, Brooks A, Chung Y, Goswami D, Anjali BA, Dewan A, Kim H, Hashizume M (2014) Tropical influenza and weather variability among children in an urban low-income population in bangladesh. Glob Health Action 7(1):24413
Karim MB, Akter MR, Haque S, Akter N, et al. (2022) Do temperature and humidity affect the transmission of sars-cov-2?-a flexible regression analysis. Annals of Data Sci 9(1):153–173
Li Q, Guan X, Wu P, Wang X, Zhou L, Tong Y, ... Feng Z (2020) Early transmission dynamics in Wuhan, China, of novel coronavirus–infected pneumonia. N Engl J Med 382:1199–1207. https://doi.org/10.1056/NEJMoA2001316
Loomis D, Richardson DB, Elliott L (2005) Poisson regression analysis of ungrouped data. Occup Environ Med 62(5):325–329
Mecenas P, Bost RDRM, Vallinoto ACR, Normando D (2020) Effects of temperature and humidity on the spread of COVID-19: a systematic review. PLoS One 15(9):e0238339. https://doi.org/10.1371/journal.pone.0238339
Nehler JA, Wedderburn RWM (1972) Generalized linear models. J R Stat Soc: Ser A (General) 135(3):370–384
Piza EL (2012) Using poisson and negative binomial regression models to measure the influence of risk on crime incident counts. Rutgers Center on Public Security
Rigby RA, Stasinopoulos DM (2005) Generalized additive models for location scale and shape. J Royal Stat Soc: Ser C (Appl Stat) 54(3):507–554
Robert WMW (1974) Quasi-likelihood functions, generalized linear models, and the gauss—newton method. Biometrika 61(3):439–447
Sun J, He W-T, Wang L, Lai A, Ji X, Zhai X, Li G, Suchard MA, Tian J, Zhou J et al (2020) Covid-19: epidemiology, evolution, and cross-disciplinary perspectives. Trends Mol Med 26 (5):483–495
The Daily Star (2020) No Need to Panic Over New Strain- Experts say; Local Virologists Found the Variant in Early November, Almost Similar to the UK one. https://www.thedailystar.net/frontpage/news/no-need-panic-over-new-strain-2016793. Online Accessed 17 Mar 2021
Wang J, Tang K, Feng K, Lv W (2020) High temperatures and high humidity reduce the transmission of COVID-19. Available at SSRN3551767
Wang Z (2011) One mixed negative binomial distribution with application. J Stat Plan Inference 141(3):1153–1160
Worldometers (2022) COVID-19 Coronavirus Pandemic. https://www.worldometers.info/coronavirus/. Online Accessed 28 Mar 2022
Xie J, Zhu Y (2020) Association between ambient temperature and COVID-19 infection in 122 cities from china. Sci Total Environ 138201:724
Yao Y, Pan J, Liu Z, Meng X, Wang W, Kan H, Wang W (2020) No association of COVID-19 transmission with temperature or UV radiation in Chinese cities. Euro Respir J 55(5):2000517. https://doi.org/10.1183/13993003.00517-2020

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.