International Journal of Statistical Sciences *ISSN 1683-5603*
Vol. 24(2) Special, December, 2024, pp 149-162
DOI: https://doi.org/10.3329/ijss.v24i20.78220

Comparing the Performance of Count Regression Models to Assess the Impact of Climate on COVID-19 Incidence in Dhaka, Bangladesh

Sahera Akter, Sabina Sharmin, Tanjina Akter and Anamul Haque Sajib[*]

Department of Statistics, University of Dhaka, Dhaka, Bangladesh
[*]Correspondence should be addressed to Anamul Haque Sajib
(Email: sajibstat@du.ac.bd)

[Received September 19, 2024; Accepted November 13, 2024]

Abstract

COVID-19 transmission has been a significant public health issue in Bangladesh since March 8, 2020. Environmental factors such as temperature, humidity, wind speed, rainfall, and visibility are thought to influence the rapid spread of the virus. This study aims to compare various count regression models to explore the relationship between these environmental factors and COVID-19 incidence. We focused on the Negative Binomial, Discrete Lindley, and Discrete Weibull regression models due to the over-dispersed nature of the COVID-19 data. Our analysis indicated that the Discrete Weibull regression model provided the best fit, as determined by AIC and dispersion values. Diagnostic plots confirmed that this model met all necessary assumptions. Additionally, a simulation study with three different scenarios was conducted to validate our findings from the real COVID-19 data. Our analysis revealed that minimum temperature and visibility are positively associated with COVID-19 transmission, while maximum temperature and humidity show a negative correlation. These insights enhance our understanding of how environmental factors impact COVID-19 outbreaks in Dhaka, offering valuable guidance for developing effective strategies to mitigate transmission in Bangladesh.

Keywords: COVID-19, Dhaka, Temperature, Humidity, Count Regression Model.

AMS Classification: 62P12, 62J99.

1. Introduction

The Coronavirus Disease 2019 (COVID-19), caused by the novel severe acute respiratory syndrome coronavirus-2 (SARS-CoV-2), has become a significant global public health challenge (Wei et al., 2022). The COVID-19 pandemic first emerged in Wuhan, China, in December 2019 and rapidly spread worldwide due to the high transmissibility and infectivity of the virus (SILVA, & VETTORE, 2024). As of September 1, 2024, the World Health Organization (WHO) has documented over 776 million confirmed cases and approximately 7.1 million deaths globally. The most frequently reported symptoms among COVID-19 patients include fever, cough, shortness of breath, and fatigue. In more severe cases, patients have presented with radiographic ground-glass lung opacities and, in some instances, progressed to acute respiratory distress syndrome (ARDS), which can prove fatal (Azuma et al., 2020).

The spread of COVID-19 is influenced by various factors, including population density, individual lifestyle habits and mobility, government restrictions, vaccination rates, and personal susceptibility. Additionally, meteorological factors such as temperature, wind speed, rainfall, humidity, sunshine duration, and visibility also play a role in shaping transmission dynamics (Zhang et al., 2024). In many Asian tropical low- and middle-income countries (LMICs), including Bangladesh, the growth rate of COVID-19 cases has been relatively slower (Haque & Rahman, 2020). Since the initial detection of cases on March 8, 2020, the WHO has reported nearly 2.1 million confirmed cases and 29,499 deaths in Bangladesh as of September 1, 2024. Despite challenges like high population density and limited testing capacity, the slower spread of COVID-19 in Bangladesh may be influenced by meteorological factors. However, the daily rise in infections and deaths continues to heighten public concern, underscoring the need to understand how Bangladesh's tropical climate may affect the spread of COVID-19. Accurately predicting these transmission patterns is essential for effective management and informed decision-making.

The literature highlights a range of meteorological factors influencing the transmission of COVID-19 worldwide. For example, Zhang et al. (2024) applied a generalized additive model (GAM) and a distributed lag nonlinear model (DLNM) to examine the relationship and lag effects between daily COVID-19 cases and meteorological factors, including temperature, relative humidity, wind speed, solar radiation, surface pressure, and precipitation in seven countries across the Americas. Their findings revealed a non-linear impact of these factors on COVID-19 transmission. Temperature exhibited two key thresholds: a positive correlation below 5°C and above 23°C, and a negative correlation in between. Similarly, relative humidity and solar radiation showed negative correlations, with slope shifts occurring around 74% humidity and 750 kJ/m² solar radiation. Lin et al. (2024) investigated how local weather conditions affect COVID-19 transmission in Taiwan using Spearman's rank correlation test. They found that new COVID-19 cases were positively linked to maximum daily temperature and relative humidity, while wind speed and diurnal temperature range were negatively related to case numbers. Ng et al. (2024) employed Pearson's and Spearman's correlation tests to analyze the relationship between weather conditions and COVID-19 cases across Malaysia. Their results indicated that average wind speed was positively correlated with confirmed COVID-19 cases, whereas average relative humidity, maximum temperature, average temperature, and minimum temperature were negatively correlated with the number of cases. Nottmeyer et al. (2023) discovered that lower temperatures and lower absolute humidity were associated with higher COVID-19 incidence. This analysis, conducted using time series analysis assuming a quasi-Poisson distribution, covered 20 different countries. Yin et al. (2022) discovered that colder, drier, and less windy conditions intensified COVID-19 transmission in Brazil using two-way fixed-effect models. Their analysis showed that daily average temperature, humidity, and wind speed negatively influenced the number of new daily cases, with humidity and temperature being the most significant factors. Khursheed et al. (2021) demonstrated that temperature, relative humidity, and absolute humidity have a significant but negative impact on the COVID-19 mortality rate in Italy, as determined using GAM and penalized spline methods. This suggests that cooler and drier conditions may facilitate virus transmission, potentially leading to higher COVID-19 death rates. Liu et al. (2020) discovered that meteorological factors independently affect COVID-19 transmission, even after accounting for population migration. Their analysis showed that low temperatures, a mild range in daily temperature, and low humidity likely promote the spread of the virus. A study in 122 cities in China, Xie and Zhu (2020), indicated that the relationship between mean temperature and confirmed COVID-19 cases was roughly linear when temperatures were below 3 °C but levelled off at temperatures above 3 °C, based on their use of generalized additive models. That means, no significant correlations were detected when the mean temperature exceeded 3°C.

Although there is extensive research on the impact of meteorological factors on COVID-19 transmission worldwide, research specifically targeting Bangladesh is relatively scarce. Miah et al. (2024) employed a multivariate generalized linear negative binomial regression model to assess the influence of climatic factors on Omicron transmission. Their analysis indicated that maximum temperature, sky clearness, wind speed, relative humidity, and air pressure all had a significant impact on the transmission of COVID-19 Omicron in Bangladesh. Parvin (2024) demonstrated a significant and positive relationship between the spread of COVID-19 and climatic factors such as temperature, humidity, rainfall, and wind speed. The Auto Regressive Distributed Lag (ARDL) model revealed that temperature and wind speed have pronounced lag effects on COVID-19 transmission in Bangladesh, while the effects of humidity and rainfall are comparatively minimal. Hasan et al. (2023) reported that both wind speed and surface pressure have a significant negative impact on COVID-19 cases and deaths. Karim and Akter (2022) found that high humidity and temperature significantly reduced the severity of COVID-19 deaths, based on their analysis using Generalized Additive Models (GAM) and Generalized Additive Models for Location, Scale, and Shape (GAMLSS). Masum and Pal (2021) used Spearman's rank correlation analysis to explore the relationship between weather-related variables and the SARS-CoV-2 outbreak in Bangladesh. Their study found a significant positive association between relative humidity and COVID-19 cases, with temperature showing mixed correlations. Humidity was also positively linked to death cases, while both rainfall and wind speed were positively associated with both the number of cases and deaths. Haque and Rahman (2020) found that, within a linear regression framework, both high temperatures and high humidity significantly lower COVID-19 transmission. This suggests that the onset of summer and the rainy season in Bangladesh could effectively reduce the spread of the virus.

A review of existing research reveals a lack of consensus on the impact of meteorological factors on COVID-19 transmission. For example, while some studies suggest a positive correlation between temperature and daily new cases (Lin et al., 2024; Parvin, 2024), others report a negative correlation (Ng et al., 2024; Yen et al., 2022; Khursheed et al., 2021). Additionally, some studies find mixed correlations (Zhang et al., 2024; Masum and Pal, 2021), and others observe no significant correlation at all (Islam et al., 2021a; Xie & Zhu, 2020). Consequently, evidence regarding the impact of weather on COVID-19 transmission remains inconclusive, likely due to significant variations in data sources, methodological approaches, study periods, socioeconomic conditions, and other factors (Wei et al., 2022). This highlights the need for more thorough and systematic research to better understand the impact of meteorological factors on the spread of COVID-19.

To the best of our knowledge, no previous study in Bangladesh has utilized the most recent meteorological data (2023). Furthermore, while most research has focused on the meteorological factors including temperature, humidity, rainfall, and wind speed, they have often neglected the role of visibility. This study addresses this gap by examining all six factors such as minimum temperature, maximum temperature, wind speed, rainfall, visibility, and humidity to understand their effects on COVID-19 transmission in Dhaka, Bangladesh. The study is innovative in its use of the latest meteorological data (2021-2023) and its application of a range of count regression models, including Negative Binomial, Discrete Lindley, and Discrete Weibull, with thorough diagnostic checks and simulation study to identify the key meteorological factors influencing COVID-19 in Bangladesh.

2. Data and Methodology

2.1 Data and Variables

Data from multiple sources were gathered for analyzing COVID-19 outbreaks in the present study. The response variable, "Daily COVID-19 Infected Cases", was collected from January 1, 2021, to November 9, 2023, using data on daily confirmed new cases and lab tests, which were sourced from the Directorate General of Health Services (DGHS) website (https://www.dghs.gov.bd). To investigate the impact of environmental factors on COVID-19 cases, daily data on minimum temperature, maximum temperature, wind speed, rainfall, visibility, and humidity were included. The daily data for minimum temperature, maximum temperature, wind speed, rainfall, and visibility were obtained from the Bangladesh Meteorological Department (BMD). Additionally, daily humidity data was sourced from the website:
https://www.timeanddate.com/weather/bangladesh/dhaka/ext.

2.2 Methodology

Given that our dependent variable, the daily number of COVID-19 cases, is a count variable exhibiting over-dispersion, we applied various classical count regression models specifically designed to handle over-dispersion to analyze the data. The candidate models used include the Negative Binomial (NB), Discrete Lindley (DL), and Discrete Weibull (DW) regression models.

Let Y_i $(i = 1,2, \dots, n)$, denote a count response variable associated with a $(p+1)$ dimensional vector of covariates, $\boldsymbol{x}_i = (1, x_{i1}, \dots, x_{ip})^T$ and $\boldsymbol{\beta} = (\beta_0, \beta_1, \dots, \beta_p)^T$ is the $(p+1) \times 1$ vector of regression coefficients corresponding to $\boldsymbol{x}_i$.

Negative Binomial Regression Model

The most widely used count regression model for analyzing over-dispersed count data is the Negative Binomial (NB) model, where $Y_i \sim \text{NB}\,(\mu_i, \nu)$, having probability mass function (pmf),

$$P(Y_i = y_i \mid \mu_i, \nu) = \frac{\Gamma(y_i + \nu)}{\Gamma(\nu)\Gamma(y_i + 1)} \left(\frac{\mu_i}{\mu_i + \nu}\right)^{y_i} \left(\frac{\nu}{\mu_i + \nu}\right)^{\nu}, \quad y_i \in \{0,1,2, \dots\},$$

where $\Gamma(\cdot)$ is the gamma function, μ_i (> 0) is the mean for the NB distribution, and ν denotes the dispersion parameter, which accounts for the additional variability. The mean can be modelled within the generalized linear model (GLM) framework as:

$$\mu_i = \exp(\boldsymbol{x}_i^T \boldsymbol{\beta}) = \exp\left(\beta_0 + \sum_{j=1}^{p} \beta_j x_{ij}\right). \tag{1}$$

Discrete Lindley Regression Model

The Discrete Lindley (DL) regression model is an adaptation of the continuous Lindley distribution, tailored specifically for count data. It is particularly useful for handling over-dispersed data, where the variance exceeds the mean, a frequent issue encountered in real-world count datasets. Suppose $Y_i \sim DL\,(\theta_i)$, which has the following form (Nguyen et al. 2023),

$$P(Y_i = y_i \mid \theta_i) = \left(1 - e^{-\theta_i}\right)^2 (1 + y_i) e^{-\theta_i y_i}, \quad y_i \in \{0, 1, 2, \dots\},\ \theta_i > 0,$$

where $$E(Y_i) = \frac{2}{e^{\theta_i - 1}} \quad \text{and} \quad \text{Var}(Y_i) = \frac{2e^{\theta_i}}{\left(e^{\theta_i - 1}\right)^2}.$$

Under GLM framework, the model can be defined as

$$\mu_i = E(Y_i \mid x_i) = \exp(\boldsymbol{x}_i^T \boldsymbol{\beta}) \quad \text{or} \quad \log(\mu_i) = \boldsymbol{x}_i^T \boldsymbol{\beta}. \tag{2}$$

where μ_i is the mean and θ_i is a shape parameter that controls the over-dispersion. A key advantage of the Discrete Lindley model is its flexibility in managing data with different levels of over-dispersion, making it a viable alternative to the NB model.

Discrete Weibull Regression Model

In real-world situations, various types of dispersion in count data can occur. To address such variability, the Discrete Weibull (DW) model is a suitable choice, as it can effectively handle any form of dispersion in count response data. The most commonly used is the Type I DW, with probability mass function,

$$P(Y_i = y_i \mid q_i, \gamma) = \begin{cases} q_i^{y_i^\gamma} - q_i^{(y_i+1)^\gamma} & for\ \ y_i = \ 0, 1, 2, \dots, \\ 0 & otherwise\ , \end{cases}$$

where $0 < q_i < 1$ and $\gamma\ (> 0)$ are the parameters with

$$E(Y_i) = \mu_i = \sum_{y_i=1}^{\infty} q_i^{y_i^\gamma} \quad \text{and} \quad \mathrm{Var}(Y_i) = 2\sum_{y_i=1}^{\infty} y_i q_i^{y_i^\gamma} - \mu_i - \mu_i^2.$$

Specifically, varying values of the shape parameter γ can accommodate different scenarios (Klakattawi et al. 2018),

• $\gamma \in (0, 1]$ represents over-dispersion, $\gamma \geq 3$ represents under-dispersion, regardless of the values of q_i

• $\gamma \in (1, 3)$ holds both over and under-dispersion depending on the values of q_i

Under GLM framework, to incorporate the effects of covariates, the Type I DW model can be written as

$$\log(-\log(q_i)) = \boldsymbol{x}_i^T \boldsymbol{\beta}, \tag{3}$$

The model that best fits our data among the three used models, will be chosen based on the lowest Akaike Information Criterion (AIC) and dispersion, with

$$\text{AIC} = 2\text{k} - 2\ \ln(\text{L}),$$

where, k is the number of regression coefficients and ln(L) is the log-likelihood of the model, and

$$\text{dispersion} = \frac{\text{Pearson}\ \chi^2}{df} = \frac{\sum_{i=1}^{n}\left(\frac{(y_i - \hat{\mu}_i)^2}{Var(\hat{\mu}_i)}\right)}{df}.$$

A model is classified as over-dispersed if the dispersion value is greater than 1, equi-dispersed if it is 1, and under-dispersed if it is less than 1.

3. Result

Table 1 provides an overview of the distribution of key variables, including climate factors. These descriptive statistics, which encompass measures such as mean, median, and standard deviation, offer insights into the overall distribution and variability of the data. This foundational overview helps in understanding the context of the data and serves as a basis for subsequent analysis and interpretation of the effects on dengue incidence.

Table 1: Summary statistics of dependent and other climatic variables from Jan 1, 2021 to Nov 9, 2023

Variable	Mean (SD)	Minimum	Maximum
COVID-19 Infected Cases	1469 (3028.36)	0	16230
Min. temperature (°C)	23.30 (4.67)	10	30.70
Max. temperature (°C)	32.07 (3.77)	15.4	40.6
Wind speed (knots)	2.49 (1.07)	0	15
Rainfall (mm)	5.08 (15.50)	0	255
Visibility (km)	4.33 (0.70)	0.4	6
Humidity (%)	74.79 (14.44)	27	100

On average, there were 1,469 COVID-19 infections per day in Dhaka, Bangladesh. Furthermore, as shown in Table 1, the variance of COVID-19 cases exceeds the mean, highlighting the over-dispersion characteristic of the count data. Additionally, summary information on the climatic factors can also be observed from Table 1. To determine if the count response exhibits any inflation, both a histogram (Figure 1) and a frequency table (Table 2) of the response were presented. Table 2 shows that during the study period, only 0.67% of people in Dhaka marked themselves safe from COVID-19 infection. Thus, overdispersion needs to be taken into account while modelling the daily COVID-19 incidence. Therefore, NB, DL, and DW regression models were fitted to analyze this data. Afterwards, to find out the most suitable model, results from goodness-of-fit tests (AIC and dispersion) were evaluated.

Table 2: Distribution of the number of daily COVID-19 cases in Dhaka city from Jan 1, 2021 to Nov 9, 2023

Distribution	Number of daily COVID-19 cases					Total
	0	1	2	3	4+	
Frequency	7	0	1	7	1028	1043
%	0.67	0	0.09	0.67	98.56	100

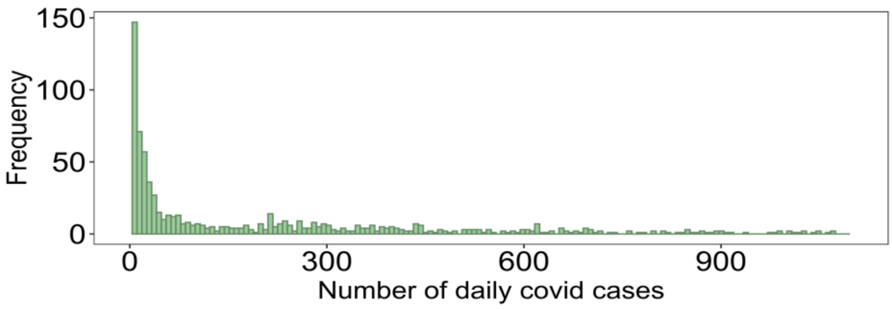


Figure 1: Histogram of daily COVID-19 cases in Dhaka city from Jan 1, 2021 to Nov 9, 2023

The AIC is a commonly used criterion for model selection, with lower AIC values indicating a better fit. Table 3 reveals that the AIC values for the NB, DL, and DW models are 15.08, 19.55, and 14.98, respectively. Thus, the DW model is preferred based on its lower AIC value. Another key criterion for selecting a count model is the dispersion parameter. For models handling over-dispersed count responses, the dispersion value should ideally be close to 1. A value near 1 suggests that the model effectively accounts for over-dispersion. Table 3 shows that, of the three models used, the DW model has a dispersion value of 0.93, which is quite close to 1.

Table 3: Goodness-of-fit tests

Model	AIC	Dispersion
NB	15.08	1.43
DL	19.55	8.54
DW	**14.98**	**0.93**

This suggests that the DW model effectively captures the over-dispersion in the data. Therefore, based on the combined result of AIC and dispersion, DW was selected as the most appropriate to analyze the over-dispersed daily COVID-19 cases.

To identify the significant climate factors affecting dengue incidence, the regression results from the Discrete Weibull model were used. Results from Table 4 show that climatic factors such as minimum temperature, maximum temperature, visibility, and humidity have a significant effect on daily COVID-19 cases, with p-values less than 0.001. However, wind speed and rainfall do not have statistically significant effects on COVID-19 incidence. The incidence risk ratio (IRR) results in Table 4 indicate that, when all predictors (minimum temperature, maximum temperature, wind speed, rainfall, visibility, and humidity) are set to zero, the baseline incidence rate of COVID-19 cases is 34,641.7 times higher. Furthermore, controlling for all other climatic factors, a 1ºC increase in minimum temperature raises the COVID-19 incidence rate by 13.7%, while a 1ºC increase in maximum temperature reduces the incidence rate by 25.8%. Additionally, for every 1 km increase in visibility, the COVID-19 incidence rate is 2.504 times higher. Similar to maximum

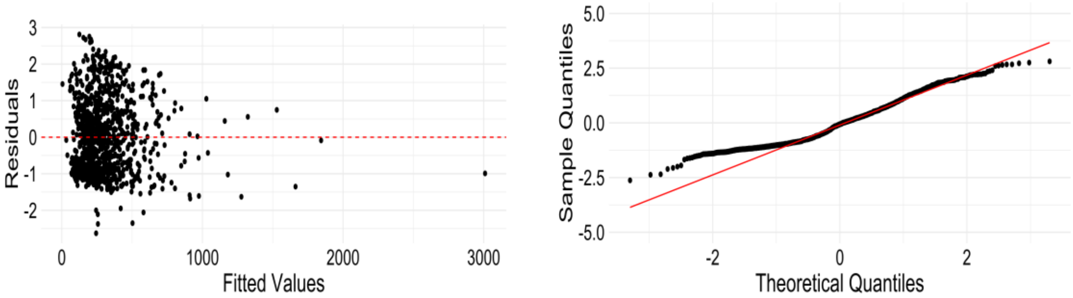
Table 4: Outputs for DW model for daily COVID-19 infected cases (Jan 1, 2021 – Nov 9, 2023)

Variable	Estimate	SE	p-value	IRR	95% CI for IRR
Intercept	10.52	0.788	<0.001***	34641.7	(7907.13,17.3e+4)
Min. temperature	0.128	0.028	<0.001***	1.137	(1.08, 1.20)
Max. temperature	-0.298	0.039	<0.001***	0.742	(0.69, 0.80)
Wind speed	0.008	0.065	0.897	1.008	(0.89, 1.15)
Rainfall	0.005	0.005	0.325	1.005	(0.99, 1.01)
Visibility	0.918	0.155	<0.001***	2.504	(1.85, 3.39)
Humidity	-0.021	0.006	<0.001***	0.979	(0.97, 0.99)

*p<0.05, **p<0.01, ***p<0.001

temperature, humidity has a negative effect on COVID-19 cases, with a 1% rise in humidity decreasing the incidence rate by 2.1%, assuming all other significant covariates are held constant.

However, it is important to assess whether the assumptions of the chosen model for analyzing daily COVID-19 cases are satisfied. To address this, model diagnostic plots are presented in Figure 2. The residuals vs. fitted values plot (a) clearly shows that the standardized residuals are approximately centred around the y = 0 line, indicating no specific pattern and suggesting randomness. In the normal q-q plot (b), the points are closely aligned with the straight line, suggesting that the model provides a good fit to the data.



(a) **(b)**

Figure 2: Model diagnostic plots for DW model (a) residuals vs. fitted plot (b) normal q-q plot

In practical applications, different models may perform optimally under specific conditions, but this does not imply that a particular model will always be the most suitable for similar scenarios. Drawing conclusions solely based on real-world observations may not be prudent. It is essential to conduct simulations using data that closely mirrors real-world conditions to verify the robustness and generalizability of the results. Thus, to verify our results from the real-life application, a simulation study was conducted in the following section.

4. Simulation Study

A simulation study was carried out to validate the findings from our real world COVID-19 incidence dataset. To achieve this, three different settings were employed to generate count responses while preserving the data composition of the COVID-19 dataset. Histograms for each simulation setting were presented to see whether the data composition of the simulated count responses closely resembled that of the daily COVID-19 incidence.

Setup 1: To generate over-dispersed count data, we assume the climatic factors as regressors and the regression parameters as, $\boldsymbol{\beta} = (\beta_0, \beta_1, \beta_2, \beta_3, \beta_4, \beta_5, \beta_6)^T$= (11.22, 0.13, -0.29, 0.02, 0.008, 0.8, -0.017) with $\nu = 0.334$. Now, generate a random sample with size n (1043) using the pmf of the NB distribution, that is for each μ_i, a corresponding y_i (i = 1, 2, … , n) was generated.

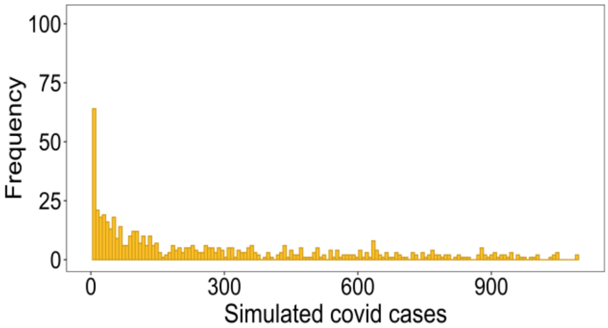
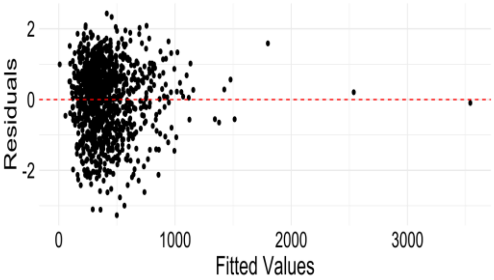


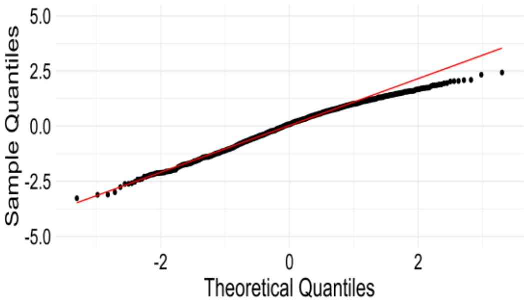
Figure 3: Histogram of the simulated count (setup 1)

Table 5: Goodness-of-fit tests

Model	AIC	Dispersion
NB	**15.45**	0.90
DL	19.22	5.23
DW	**15.51**	**1.08**



(a)



(b)

Figure 4: Model diagnostic plots for DW model (a) residuals vs. fitted plot (b) normal q-q plot under setup 1

Setup 2: Say, the regression parameters are $\boldsymbol{\beta} = (\beta_0, \beta_1, \beta_2, \beta_3, \beta_4, \beta_5, \beta_6)^T$= (11.20, 0.13, -0.29, 0.02, 0.008, 0.8, -0.02). Now, generate a random sample with size n (1043) using the pmf of the DL distribution, that is for each μ_i, there was a corresponding y_i (i = 1, 2, … , n).

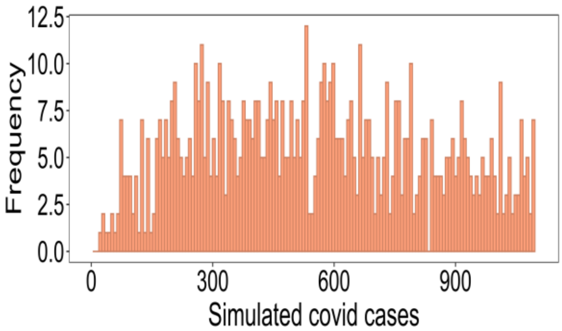
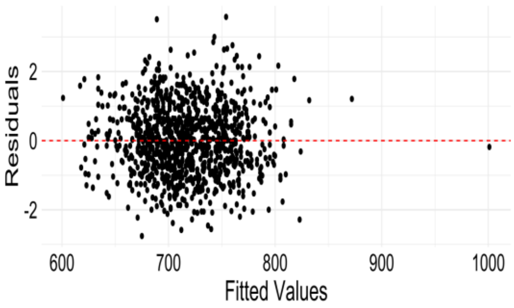


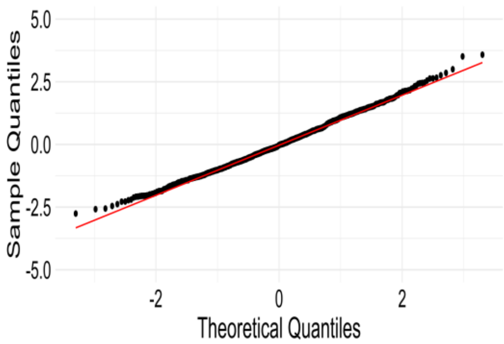
Figure 5: Histogram of the simulated count (setup 2)

Table 6: Goodness-of-fit tests

Model	AIC	Dispersion
NB	15.18	0.96
DL	15.17	0.93
DW	15.18	**0.98**



(a)



(b)

Figure 6: Model diagnostic plots for DW model (a) residuals vs. fitted plot (b) normal q-q plot under setup 2

Setup 3: Suppose the regression parameters for data generation are $\boldsymbol{\beta} = (\beta_0, \beta_1, \beta_2, \beta_3, \beta_4, \beta_5, \beta_6)^T$= (10.45, 0.13, -0.29, 0.02, 0.006, 0.92, -0.021) with γ = 0.472. Now, generate a random sample with size n (1043) using the pmf of the Type I DW distribution, that is for each q_i, there will be a corresponding y_i (i = 1, 2, … , n).

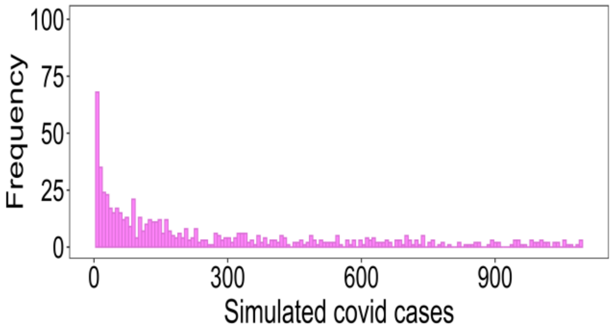
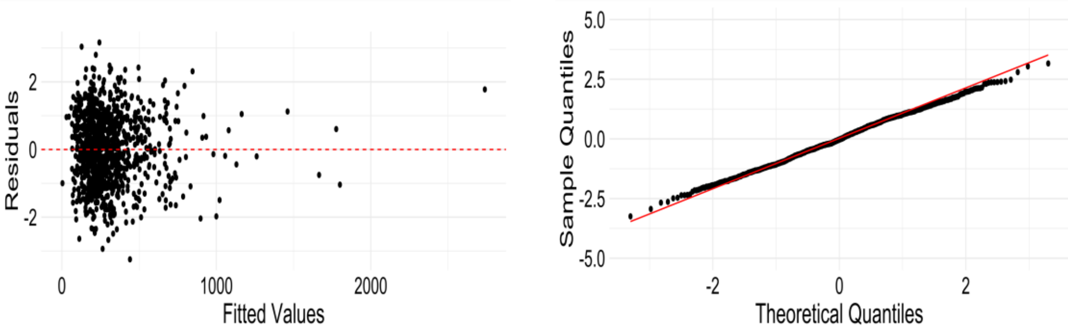


Figure 7: Histogram of the simulated count (setup 3)

Table 7: Goodness-of-fit tests

Model	AIC	Dispersion
NB	14.95	1.52
DL	19.07	9.12
DW	**14.93**	**1.01**



(a) **(b)**

Figure 8: Model diagnostic plots for DW model (a) residuals vs. fitted plot (b) normal q-q plot under setup 3

To verify the data composition, the histograms of the simulated count responses (Figures 3, 5, and 7) were examined across three setups, revealing signs of over-dispersion. Afterwards, the NB, DL, and DW regression models were fitted on the simulated data for each setup. Based on the model selection criteria (AIC and dispersion), DW was selected as the best model in each setup, which was also supported by the model diagnostic plots (Figures 4, 6, and 8).

In keeping with the real-life data composition, the simulation study generated count responses across three different setups. In all scenarios, the results were clearly demonstrated through model selection and diagnostic plots, effectively validating the real data application of COVID-19 infected cases.

5. Discussion

The emergence of the novel SARS-CoV-2 virus has posed a significant threat to global health, leading to the COVID-19 pandemic since December 2019 (Hossain et al., 2024). This study investigates the impact of climatic factors on daily COVID-19 cases in Dhaka, Bangladesh, over a three-year period from 2021 to 2023. Due to the significant fluctuations in daily case counts, several count regression models were explored, with adjustments made to address over-dispersion.

Among them, the discrete Weibull regression model was identified as the best fit for the data, based on the AIC, dispersion, and graphical assessment.

The results of this study indicate that certain meteorological factors, including minimum temperature, maximum temperature, visibility, and humidity, significantly influence the spread of COVID-19. In contrast, wind speed and rainfall showed no impact on transmission rates. The analysis also suggests a potential nonlinear relationship between temperature and daily COVID-19 cases. Specifically, a minimum temperature range of 10°C to 30.7°C was positively correlated with daily case counts, whereas maximum temperatures range between 15.4°C and 40.6°C exhibited an inverse association. These results are in line with research by Liu et al. (2022) using meteorological and epidemiological data from 153 countries and Islam et al. (2021b) in Bangladesh. However, previous studies have produced mixed findings on this topic. For instance, Sobral et al. (2020) and Wu et al. (2020) observed a negative correlation between temperature and COVID-19 transmission, regardless of whether the temperature was high or low. Conversely, Xie and Zhu (2020) found that daily COVID-19 cases increased by 4.86% (95% CI: 3.21%, 6.51%) for every 1°C rise in temperature when ambient temperatures were below 3°C in a study of 122 Chinese cities. Although the findings in this study may not fully replicate those of previous research, they offer partial confirmation and useful insights. One plausible explanation is that higher temperatures may reduce the viability of SARS-CoV-2, while colder temperatures, particularly in winter, may weaken human innate immunity, as noted by Wu et al. (2020).

According to our study findings, wind speed has no significant effect on COVID-19 outbreaks. The finding is consistent with the study in Ukraine (Podavalenko, 2023), Bangladesh (Islam et al., 2021a) and New York (Bashir et al., 2020). Contrary to these results, a significant correlation between COVID-19 and wind speed was reported in Bangladesh (Parvin, 2024) and Taiwan (Lin et al., 2024). The results of this study also suggest that rainfall is an insignificant risk factor for COVID-19 transmission, consistent with the findings of Zhang et al. (2024) in seven countries across the Americas, Nawi et al. (2022) in Malaysia, Rendana (2020) in Indonesia, and Bashir et al. (2020) in New York. However, studies by Masum and Pal (2021) and Sobral et al. (2020) reported contradictory findings. This discrepancy could be attributed to the highly variable dynamics of COVID-19 outbreaks across different countries and regions, as noted by Dong et al. (2020).

Visibility, defined as the clarity of the atmosphere or the distance at which objects can be clearly seen, has emerged as a statistically significant environmental factor for COVID-19 cases, contributing to an increase in transmission. This may be because clear skies encourage more outdoor gatherings, leading to greater human interaction and a higher potential for transmission, especially when social distancing and mask usage are not adequately maintained. Additionally, the study found that humid weather reduces the spread of COVID-19, as it has a significant negative impact on transmission rates. This finding aligns with previous studies conducted in Bangladesh (Karim and Akter, 2022; Haque and Rahman, 2020) and Malaysia (Ng et al., 2024). However, contradictory results were reported in studies by Lin et al. (2024) and Rendana (2020). One explanation for the protective effect of humidity is that higher moisture levels help keep the nasal and throat membranes moist, which aids in capturing dirt, bacteria, and viruses before they reach the lungs (Haque and Rahman, 2020).

The findings were validated using simulation studies across three distinct setups. These simulations demonstrated that the discrete Weibull regression model provided the best fit for analyzing the over-dispersed daily COVID-19 case data.

6. Conclusion

This study explores the relationship between climate factors and COVID-19 incidence in Dhaka, Bangladesh. To analyze the data, NB, DL, and DW regression models were fitted, with the DW model being selected as the most appropriate based on AIC and dispersion criteria. Additionally, model diagnostic plots were examined to ensure the assumptions of the chosen model were met. The findings from the real data analysis were further validated through a simulation study. Our research confirms a significant correlation between COVID-19 transmission and environmental factors, particularly minimum temperature, maximum temperature, visibility, and humidity. Minimum temperature and visibility show a positive influence on COVID-19 cases, while maximum temperature and humidity have a negative impact. However, rainfall and wind speed do not appear to influence transmission in this study. These insights can aid policymakers in developing a climate-based warning system. Future efforts may involve constructing a comprehensive model that incorporates immunological, entomological, demographic, and climatic data to more accurately predict COVID-19 cases.

However, this study is not without limitations. First, relying solely on COVID-19 data from Dhaka may not reflect conditions in other areas with different environmental and social contexts. Additionally, incomplete or inconsistent data such as case underreporting and variability in weather station readings could affect findings. While six meteorological factors (minimum temperature, maximum temperature, wind speed, rainfall, visibility, and humidity) were analyzed, these factors may interact with other variables not included in the study, such as population density, transportation patterns, healthcare access, and socioeconomic conditions. Moreover, the study did not account for seasonality, which could reveal critical transmission patterns. Future research should incorporate broader regional data and seasonal trends to support more targeted COVID-19 policies.

Reference

[1] Azuma, K., Yanagi, U., Kagi, N., Kim, H., Ogata, M., and Hayashi, M. (2020). Environmental factors involved in SARS-CoV-2 transmission: effect and role of indoor environmental quality in the strategy for COVID-19 infection control. Environmental health and preventive medicine, 25, 1-16.

[2] Bashir, M. F., Ma, B., Komal, B., Bashir, M. A., Tan, D., and Bashir, M. (2020). Correlation between climate indicators and COVID-19 pandemic in New York, USA. Science of the Total Environment, 728, 138835.

[3] Coccia, M. (2020). Diffusion of COVID-19 outbreaks: the interaction between air pollution-to-human and human-to-human transmission dynamics in hinterland regions with cold weather and low average wind speed.

[4] Dong, E., Du, H., and Gardner, L. (2020). An interactive web-based dashboard to track COVID-19 in real time. The Lancet infectious diseases, 20(5), 533-534.

[5] Haque, S. E., & Rahman, M. (2020). Association between temperature, humidity, and COVID-19 outbreaks in Bangladesh. Environmental science & policy, 114, 253-255.

[6] Hasan, M. N., Islam, M. A., Sangkham, S., Werkneh, A. A., Hossen, F., Haque, M. A., ... and Sarkodie, S. A. (2023). Insight into vaccination and meteorological factors on daily COVID-19 cases and mortality in Bangladesh. Groundwater for Sustainable Development, 21, 100932.

[7] Hossain, M., Mannan, R., Islam, S., Banu, L. A., Jamee, A. R., Hassan, Z., ... and Azad Khan, A. K. (2024). Unveiling the occurrence of COVID-19 in a diverse Bangladeshi population during the pandemic. Frontiers in Public Health, 12, 1363971.

[8] Islam, A. R. M. T., Hasanuzzaman, M., Azad, M. A. K., Salam, R., Toshi, F. Z., Khan, M. S. I., ... and Ibrahim, S. M. (2021b). Effect of meteorological factors on COVID-19 cases in Bangladesh. Environment, Development and Sustainability, 23, 9139-9162.

[9] Islam, M. S., Rahman, M., Tusher, T. R., Roy, S., and Razi, M. A. (2021a). Assessing the relationship between COVID-19, air quality, and meteorological variables: a case study of Dhaka City in Bangladesh. Aerosol and Air Quality Research, 21(6), 200609.

[10] Karim, R., and Akter, N. (2022). Effects of climate variables on the COVID-19 mortality in Bangladesh. Theoretical and Applied Climatology, 150(3), 1463-1475.

[11] Khursheed, A., Mustafa, F., and Akhtar, A. (2021). Investigating the roles of meteorological factors in COVID-19 transmission in Northern Italy. Environmental Science and Pollution Research, 28(35), 48459-48470.

[12] Klakattawi, H. S., Vinciotti, V., and Yu, K. (2018). A simple and adaptive dispersion regression model for count data. Entropy, 20(2), 142.

[13] Lin, W. Y., Lin, H. H., Chang, S. A., Chen Wang, T. C., Chen, J. C., and Chen, Y. S. (2024). Do Weather Conditions Still Have an Impact on the COVID-19 Pandemic? An Observation of the Mid-2022 COVID-19 Peak in Taiwan. Microorganisms, 12(5), 947.

[14] Liu, J., Zhou, J., Yao, J., Zhang, X., Li, L., Xu, X., ... and Zhang, K. (2020). Impact of meteorological factors on the COVID-19 transmission: A multi-city study in China. Science of the total environment, 726, 138513.

[15] Liu, M., Li, Z., Liu, M., Zhu, Y., Liu, Y., Kuetche, M. W. N., ... and Tao, L. (2022). Association between temperature and COVID-19 transmission in 153 countries. Environmental Science and Pollution Research, 1-11.

[16] Masum, M. H., and Pal, S. K. (2021). Association between Meteorological Parameters and SARS-CoV-2 Outbreak Trends in Major Divisions of Bangladesh. Journal of Scientific Research, 13(3).

[17] Miah, M. M., Faruk, M. O., Pingki, F. H., and Al Neyma, M. (2024). The effects of meteorological factors on the COVID-19 omicron variant in Bangladesh. International Journal of Environmental Health Research, 34(1), 514-525.

[18] Nawi, N. F. A. M., Dom, N. C., Dapari, R., and Precha, H. S. A. (2022). Evaluation of The Effect of Rainfall on COVID-19 Transmission In Terengganu, Malaysia. Journal of Sustainability Science and Management, 17(12), 2-12.

[19] Ng, K. Y., Hasan, M. Z., and Rahman, A. (2024). Examining the roles of meteorological variables in COVID-19 spread in Malaysia. Aerobiologia, 1-16.

[20] Nguyen, M., Nguyen, M. V., and Le, N. T. (2023). Using the Discrete Lindley Distribution to Deal with Over-dispersion in Count Data. Austrian Journal of Statistics, 52(3), 96-113.

[21] Nottmeyer, L., Armstrong, B., Lowe, R., Abbott, S., Meakin, S., O'Reilly, K. M., ... and Sera, F. (2023). The association of COVID-19 incidence with temperature, humidity, and UV radiation–A global multi-city analysis. Science of The Total Environment, 854, 158636.

[22] Parvin, R. (2024). COVID-19 and the impact of climatic parameters: a case study of Bangladesh. Theoretical and Applied Climatology, 155(1), 645-659.

[23] Podavalenko, A. P., Malysh, N. H., Kuzmenko, O. V., Zadorozhna, V. I., Kolomiiets, S. V., and Chemych, O. M. (2023). Influence of meteorological factors on COVID-19 incidence in the conditions of Ukraine.

[24] Rendana, M. (2020). Impact of the wind conditions on COVID-19 pandemic: a new insight for direction of the spread of the virus. Urban climate, 34, 100680.

[25] SILVA, A. N. D., and VETTORE, M. V. (2024). Factors associated with dental students' mental health during the COVID-19 pandemic: a longitudinal study. Brazilian Oral Research, 38, e090.

[26] Sobral, M. F. F., Duarte, G. B., da Penha Sobral, A. I. G., Marinho, M. L. M., and de Souza Melo, A. (2020). Association between climate variables and global transmission of SARS-CoV-2. Science of The Total Environment, 729, 138997.

[27] Wei, Y., Dong, Z., Fan, W., Xu, K., Tang, S., Wang, Y., and Wu, F. (2022). A narrative review on the role of temperature and humidity in COVID-19: Transmission, persistence, and epidemiological evidence. Eco-Environment & Health, 1(2), 73-85.

[28] World Health Organization. (2023). WHO Coronavirus (COVID-19) dashboard. Retrieved September 1, 2024, from https://data.who.int/dashboards/covid19

[29] Wu, Y., Jing, W., Liu, J., Ma, Q., Yuan, J., Wang, Y., ... and Liu, M. (2020). Effects of temperature and humidity on the daily new cases and new deaths of COVID-19 in 166 countries. Science of the Total Environment, 729, 139051.

[30] Xie, J., and Zhu, Y. (2020). Association between ambient temperature and COVID-19 infection in 122 cities from China. Science of the Total Environment, 724, 138201.

[31] Yin, C., Zhao, W., and Pereira, P. (2022). Meteorological factors' effects on COVID-19 show seasonality and spatiality in Brazil. Environmental research, 208, 112690.

[32] Zhang, H., Wang, J., Liang, Z., and Wu, Y. (2024). Non-linear effects of meteorological factors on COVID-19: An analysis of 440 counties in the Americas. Heliyon, 10(10).