E-ISSN 2586-9027
Homepage : *https://tci-thaijo.org/index.php/SciTechAsia* Science & Technology Asia
Vol.30 No.2 April - June 2025 Page: [108-124]

*Original research article*

# Identifying the Effects of Environmental Factors on Dengue Transmission in Dhaka, Bangladesh: A Parametric Count Regression Approach

Anamul Haque Sajib[1,*], Sahera Akter[1], Faysal A Chowdhury[2]

[1] *Department of Statistics, University of Dhaka, Dhaka-1000, Bangladesh*
[2] *Florida Gulf Coast University, Department of Mathematics, Fort Myers, Florida 33965, United States*

Received 3 October 2024; Received in revised form 19 May 2025
Accepted 4 June 2025; Available online 18 June 2025

## ABSTRACT

Bangladesh faces ongoing public health challenges from dengue fever, particularly in urban areas like Dhaka. The spread of this mosquito-borne disease is influenced by climate, with temperature, rainfall, humidity, and wind speed affecting its occurrence and distribution. This study investigates the relationship between these environmental factors and dengue transmission in Dhaka from 2021 to 2023. Various count models—Poisson, Negative Binomial, discrete Lindley and Weibull, zero-inflated, and hurdle models—were applied to analyze dengue incidence. Model selection was based on AIC, dispersion, and predictive criteria. A simulation study supported the findings, consistently identifying the discrete Weibull model as the best fit. Results showed that maximum temperature and wind speed were negatively associated with dengue cases, while minimum temperature and humidity had a positive effect. Rainfall and visibility showed no significant impact. This study enhances understanding of how climate influences dengue in Dhaka and supports the development of effective prevention strategies, including a potential climate-based warning system for Bangladesh.

**Keywords:** Dengue; Dhaka; Humidity; Incidence; Temperature

## 1. Introduction and Preliminaries

Dengue, a disease transmitted primarily by the biting of infected Aedes mosquitoes, predominantly Aedes aegypti and Aedes albopictus species, has emerged as a significant global public health concern due to its widespread distribution, in-

***Corresponding author:** sajibstat@du.ac.bd **doi:** 10.14456/scitechasia.2025.29

creasing incidence, and substantial socioeconomic burden. Dengue Hemorrhagic Fever (DHF) and Dengue Shock Syndrome (DSS) are more severe than the other type of dengue, common dengue fever. Especially in tropical and subtropical regions, it is regarded as an alarming arbovirus comprising four different strains: DENV-1, DENV-2, DENV-3, and DENV-4, that cause dengue virus (DENV), of which DENV-3 and DENV-4 serotypes are typically regarded as belonging to the virulent group.

Dengue has become a global problem, striking tropical and subtropical countries alarmingly frequently since its first known outbreak in the 1950s. Over 100 countries have an endemic case of dengue, which is primarily found in tropical and subtropical regions of the world [1]. Around 75% of dengue cases worldwide occur in Southeast Asia and the Western Pacific region, which also has a heavy disease burden [2]. According to estimates from the World Health Organization (WHO), about 390 million cases of dengue fever occur each year, putting almost half of the world's population at risk of contracting the virus [3].

Bangladesh has become a conducive environment for the transmission of the dengue virus with its tropical climate. The initial recorded outbreak of Dengue in Bangladesh occurred in 1964 within Dhaka city, with subsequent sporadic cases of Dengue Fever reported during the periods of 1977-1978 and 1996-1997 [4]. But midway through 2000, there was a dengue pandemic that spread to all of the main cities and resulted in 5521 officially documented cases and 93 fatalities [5, 6].

From 2000 to 2009, Dhaka held the status of being the most endemic metropolitan area in Bangladesh, with 91% of all reported dengue cases recorded within its borders [7]. The outbreak in 2019 stands as the nation's most extensive to date, surpassing the 2018 outbreak, which was the second largest. In 2019, there were a staggering 87,953 reported cases of dengue, resulting in 81 deaths, marking a nearly ninefold rise compared to the previous record of 10,148 cases in 2018 [8]. Dengue has increasingly become a pressing issue in Bangladesh, exacting a significant human toll each year. From January to December 2022, the country witnessed 59,196 cases of dengue and 258 fatalities [9].

Numerous studies have been conducted to investigate the relationship between various climatic variables and dengue incidence worldwide. Research results from a study indicated that there is a strong association between serotype DENV-3 and a heightened occurrence of acute dengue cases [10]. A separate study employed Poisson and negative binomial regression models to estimate the effect of different predictor variables on the dengue [11]. Other researchers undertook a study to explore the association between climate variables and dengue fever occurrences in Dhaka, Bangladesh, analyzing data spanning from 2013 to 2020 [5]. They fitted Poisson, zero-inflated Poisson, and negative binomial regression models, and their findings revealed a positive correlation between dengue outbreaks and maximum, minimum temperature, humidity, and wind speed, with rainfall and sunshine hours showing a negative association. The results imply that temperature significantly influences the transmission of dengue fever, indicating that mosquito vectors tend to be more active in warmer conditions.

Additionally, a study on daily dengue incidence has been conducted that used data from 2021 to 2022 and employed different count models, including zero-inflated Conway-Maxwell Poisson regression, zero-inflated Poisson regression, zero-truncated Conway-Maxwell Poisson regression, and zero-truncated negative binomial [12]. Different potential environmental factors for dengue incidence have been determined. As per their study, humidity, precipitation, and wind speed were all found to have a significant effect on dengue fever. A separate study focused on the dengue situation in Dhaka, Bangladesh, by modeling data from 2000 to 2010 [13]. In their research, they used the Poisson time series model combined with the DLM (distributed lag model) and found effects of temperature and humidity on dengue outbreaks but no effect of rainfall.

Additionally, research revealed significant association between dengue cases and climate factors, analyzing 9 years (2014-2022) divisional data for model performance analysis [14]. Furthermore, a study modeled data over a decade (2000–2009) from Dhaka, Bangladesh, using a negative binomial generalized linear model and found a positive impact of temperature and rainfall but no impact of humidity [15]. An investigation conducted in Kolkata, India, constructed a regression model aimed at predicting the occurrence of dengue fever [16]. They employed data on dengue incidence along with climate metrics gathered between 2005 and 2016. The study outcomes unveiled a notable association between instances of dengue fever and both maximum and minimum temperatures, humidity, and rainfall. The findings suggested that elevated levels of maximum and minimum temperatures, humidity, and rainfall were positively associated with heightened dengue case counts.

A study in Pakistan discovered a robust relationship between rainfall and dengue fever occurrences in Lahore, Pakistan, spanning from 2013 to 2014 [17]. Furthermore, the research revealed correlations between temperature, humidity, and the spread of dengue in the region. Rainfall and the highest temperature were shown to be substantially correlated with dengue incidence in a separate study on the spatial-temporal effect of dengue disease in Pakistan from 2006 to 2017 [18]. In order to examine the relationship between climatic factors and dengue transmission, the study combined data from several sources, such as weather reports, demographic statistics, and dengue case records. The study's findings demonstrated a positive correlation between dengue incidence and greater maximum temperatures and rainfall.

Examining the connection between climate variables and dengue fever in Nepal spanning from 2010 to 2017, a study revealed a strong correlation between dengue fever and minimum and maximum temperatures as well as humidity [19]. The findings indicated that increased minimum and maximum temperatures, along with higher humidity, were linked to elevated dengue incidence. However, the study did not identify a significant relationship between rainfall and dengue occurrence.

A study in Malaysia found a positive correlation between temperature and rainfall, with no effect of precipitation on dengue [20]. Additionally, a study analyzed data spanning a five-year period (2008-2012) from China [21]. They employed a Poisson regression model and found that humidity and precipitation had a significant impact, while there was no observed effect of temperature on dengue incidence.

In Sulawesi, Indonesia, researchers investigated the connection between climate variables and dengue disease by developing a model for dengue prediction in the future [22]. All of this data was gathered throughout 2010 and 2015. They found that while there were negative associations between dengue fever and rainfall and humidity, there was a positive correlation between dengue fever and mean temperature.

Similarly, a study investigated the impact of climate factors on dengue fever in Davao, Philippines, analyzing meteorological and dengue data from 2011 to 2015 [23]. They identified low temperatures and moderate rainfall as significant contributors to dengue incidence.

A study in Itan fitted the ARIMA model with the transfer function on dengue and climatic data from 2003 to 2009 and found the opposite effects of temperature (positive) and humidity (negative) on dengue fever, whereas there was no significance of precipitation [24]. Another study considered temperature and rainfall as covariates and found these to be influential environmental factors for dengue outbreaks by analyzing data from a period of 8 years (1998–2015) [25]. Unique results were seen in a separate study, where all the climate variables, such as temperature, humidity, and rainfall, were statistically negatively associated with dengue cases [26]. They utilized data on Rio Branco, Brazil, from 2001 to 2012 and analyzed it using generalized autoregressive moving average models with a negative binomial distribution.

While numerous researchers have sought to ascertain the prevalence of dengue in Bangladesh, to the best of our knowledge, no study has yet been carried out in Bangladesh where comprehensive model comparison was made as well as up-to-date meteorological data (2023) having been utilized. Additionally, the majority of studies have attempted to delineate the correlation between temperature, humidity, rainfall, and wind speed but have often overlooked the influence of visibility.

Thus, this study has taken into account all six factors such as, minimum temperature, maximum temperature, wind speed, rainfall, visibility, and humidity in order to comprehend their impact on the transmission of dengue in Dhaka, Bangladesh. In this study, we thoroughly examine a comprehensive range of count regression models, covering nearly all possible variations, which has not been done in previous research.

Alongside other established count regression models, the discrete Weibull model has recently gained attention in the literature. A recent study on COVID-19 demonstrated the effectiveness of the discrete Weibull model, showing that it outperformed several other count regression models in accurately capturing the dynamics of COVID-19 count data [27]. This finding primarily justifies the use of the discrete Weibull model in our study. The novelty of this study lies in the utilization of the latest meteorological data (2021–2023) and the application of almost all possible count regression models to determine the influential climatic factors affecting dengue incidence in Bangladesh.

## 2. Materials and Methods

### 2.1 Data and variables

In this study, data from several sources have been extracted for conducting analysis on dengue outbreaks. To carry out the study, data on the response variable, “Daily Dengue Infected Cases” from 2021 to 2023, was gathered from the daily reports of the Directorate General of Health Services (DGHS), Dhaka, Bangladesh’s Health Emergency Operation Center & Control Room, which can be accessed at http://www.dghs.gov.bd. To explore the influential environmental factors on dengue cases, minimum temperature, maximum temperature, wind speed, rainfall, visibility, and humidity are considered as independent variables based on extensive literature. The daily minimum temperature, maximum temperature, wind speed, rainfall, and visibility data have been collected from Bangladesh Meteorological Department. Further, data on daily humidity has been extracted from the website https://www.timeanddate.com/weather/bangladesh/dhaka/ext. A negligible number of cases (about 0.3%) were missing, and after omitting them, 1,092 observations remained.

### 2.2 Modeling count data

Count data refer to the number of times an event occurs within a fixed period of time. For example, the number of dengue infected cases in a particular month, the number of deaths during a year, the number of abusive acts a child aged 5-14 years was exposed to during a month, and so on. To handle count data, there are various distributions available, the most well-known of which is the Poisson distribution. However, the equi-dispersion assumption of Poisson distribution, that is, when variance is equal to its mean, limits its applicability in numerous real-world scenarios.

Count data frequently show up in situations where excess variation between counts are seen in practice, yielding over-dispersion (variance > mean). One of the most widely used distributions to manage these kinds of over-dispersed count data is the Negative Binomial (NB), which is used when there is more count fluctuation than expected. Other distributions, such as the discrete Lindley, and discrete Pareto, etc., have also been occasionally considered as a choice in the presence of over-dispersion. A flexible choice to handle any kind of dispersed data can be the Generalized Poisson, discrete Weibull, etc. Moreover, in some applications, excess or absence of any particular count(s) are observed as a result of mixture model. That is, when the counts come from two different portions of the population, using only the classical count model may result in misleading inferences. These situations can be handled by the two-component models. For instance, to deal with excess zeros in the count data, the zero-inflated and hurdle models have been used. All the mentioned models can be generalized to linear models through which the relationship between the predictors and the count of occurrences can be explained.

Let $Y_i (i = 1, 2, \ldots, n)$, denote a count response variable associated with a $(p + 1)$ dimensional vector of covariates, $x_i = (1, x_i 1, \ldots, x_i p)^T$ and $\beta = (\beta_0, \beta_1, \ldots, \beta_p)^T$ is the $(p + 1) \times 1$ vector of regression coefficients corresponding to $x_i$.

The Poisson regression model is one of the most popular choices to model any count data with probability density function.

$$P(Y_i = y_i|\mu_i) = (e^{-\mu_i} \mu_i^{y_i})/(y_i!), \quad y_i \in \{0, 1, 2, ...\}, \tag{2.1}$$

where the mean and variance of the Poisson variate are given by $\mu_i$ (>0), which is the fundamental property of the Poisson distribution, called “equi-dispersed” [28]. Under the GLM setup, this distribution is generalized by relating $\mu_i$ to a set of covariates $x_i = (1, x_{i1}, x_{i2}, ..., x_{ip})^T$ with corresponding parameters $\beta$, via the log link function,

$$\log(\mu_i) = x_i^T \beta = \beta_0 + \sum_{j=1}^{p} \beta_j x_{ij}. \tag{2.2}$$

In model (2.2), the parameter $\beta_j$ can be interpreted as the expected change in the log of the mean per unit change in the predictor $x_{ij}$ [29]. Although the Poisson model is conventionally considered to be the most basic model for analyzing counts, certain departures from the assumptions of the Poisson distribution led researchers to opt for alternative specifications of the Poisson model, which may allow over-dispersion or under-dispersion or over-abundance or under-abundance of a particular count, etc.

The Negative Binomial (NB) model could be a good choice to handle the issue of over-dispersion. This is the generalization of the Poisson regression that relaxes the restriction of the variance being equal to the mean by incorporating a dispersion parameter. One can assume that $Y_i \tilde{} NB(\mu_i, \nu)$ with probability density function,

$$P(Y_i = y_i|\mu_i, \nu) = \frac{\Gamma(y_i + \nu)}{\Gamma(\nu)\Gamma(y_i + 1))} (\frac{\mu_i}{\mu_i + \nu})^{y_i} (\frac{\nu}{\mu_i + \nu})^{\nu}, \quad y_i \in \{0, 1, 2, \ldots\}, \quad (2.3)$$

where $\Gamma(.)$ is the gamma function, $\mu_i$ (>0) is the mean for the NB distribution, and $\nu$ denotes the dispersion parameter, which allows for considering the variability. The mean can be modeled within the GLM framework as:

$$\mu_i = exp(x_i^T \beta) = exp(\beta_0 + \sum_{j=1}^{p} \beta_j x_{ij}), \quad (2.4)$$

and the variance is given by $Var(Y_i) = \frac{\mu_i + \mu_i^2}{\nu}$. Variance for both the Poisson and NB will be identical when $\nu \to \infty$. The discrete Lindley (DL) model, which is a discretization of the continuous Lindley model, has gained popularity recently for having the capacity to handle actuarial data that frequently exhibits the over-dispersion phenomenon more flexibly than the Poisson distribution [30]. Suppose that we have a sample of n observations $y_1, y_2, \ldots, y_i$ are generated from discrete Lindley distribution with $Y_i \tilde{} DL(\theta_i)$, which has the following form

$$P(Y_i = y_i) = (1 - e^{-\theta_i})^2 (1 + y_i) e^{-i y_i}, \quad y_i \in \{0, 1, 2, \ldots\}, \theta_i > 0. \quad (2.5)$$

It can be shown that [31],

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

While incorporating the effect of the covariates, this model can be generalized as a linear model such as

$$log(\mu_i) = x_i^T \beta. \quad (2.6)$$

Even though the given models are suitable for over-dispersion relative to Poisson, they provide misleading results in cases of under-dispersed count data. Therefore, reasonable models for fitting under-dispersed counts must be taken into consideration. When count data exhibit under-dispersion, the Negative Binomial regression model is not suitable. A different regression model that can handle under-dispersion needs to be introduced. The discrete Weibull (DW) model is one of those models that can handle over-dispersion, under-dispersion, or both. The most common type in the literature is the first type of DW model, which has the following form.

$$P(Y_i = y_i|q_i, \alpha) = \begin{cases} q_i^{y_i^{\alpha}} - q_i^{(y_i+1)^{\alpha}} & \text{for } y_i = 0, 1, 2, \ldots, \\ 0 & \text{Otherwise,} \end{cases} \quad (2.7)$$

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

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

In particular, different values of the shape parameter, $\alpha$ may allow different cases, for instance, $\alpha \in (0, 1]$ represents over-dispersion, $\alpha \geq 3$ represents under-dispersion, regardless of the values of $q_i$, and $\alpha \in (1, 3)$ holds both over and under-dispersion depending on the values of $q_i$ [32]. Formally, the effects of the covariates can be incorporated through the link function,

$$\log(-\log(q_i)) = x_i^T \beta, \qquad (2.8)$$

which is the desired generalization of the Type I DW regression model. Real-world situations may involve a excess number of zeros in count data, such as the number of dengue infected cases in a given month, the number of deaths in an accident, etc. In this situation, usually one may face the problem of over-dispersion, but modeling counts using the classical methods may led to inappropriate statistical inference. Hence, further modification is needed to address this issue. Particularly, the zero-inflated and hurdle models were recommended for zero inflation. The zero-inflated models consist of two parts, one for zero counts and the other one for positive counts. Thus, one obtains the two-component mixture models, as

$$P(Y_i = y_i) = \begin{cases} \pi_i + (1-\pi_i) f_P(0) & \text{for } y = 0, \\ (1-\pi_i) f_P(y_i) & \text{for } y \in \{1, 2, \ldots\}, \end{cases} \qquad (2.9)$$

where $f_P(\cdot)$ is the pdf of the parent count model, which can be either Poisson or Negative Binomial, and $\pi_i = P(C_i = 0)$ denotes the zero-inflation parameter ($0 < \pi_i < 1$) with $C_i$ denoting the class membership, given by

$$C_i = \begin{cases} 0 & \text{for } y = 0, \\ 1 & \text{for } y \in \{1, 2, \ldots\}, \end{cases}$$

Following that, the two predictors build the relationship between responses and covariates

$$\log(\mu_i) = x_i^T \beta, \text{ and}$$
$$\text{logit}(\pi_i) = x_i^T \gamma, \qquad (2.10)$$

where $\gamma = (\gamma_0, \gamma_1, \ldots, \gamma_p)^T$ is a another set of real-valued coefficients. Hurdle models can be used as an alternative to handle the issue of zero inflation [33]. But in cases in which all zero counts arise from the structural zeros, the hurdle model is superior to the zero-inflated model.This model combines a binary process, π, and a truncated-at-zero count model, which has the form:

$$P(Y_i = y_i) = \begin{cases} \pi_i & \text{for } y = 0, \\ (1-\pi_i) \frac{f_P(y_i)}{1 - f_P(0|x_i)} & \text{for } y \in \{1, 2, \ldots, n\}, \end{cases} \qquad (2.11)$$

where $\pi_i$ determines the binary choice between zero and a positive count and $f_P(\cdot)$ refers to the parent process, which can be Poisson or NB model. Now the link between responses and covariates can be expressed as

$$\log(\mu_i) = x_i^T \beta, \text{ and}$$
$$\text{logit}(\pi_i) = x_i^T \gamma. \qquad (2.12)$$

### 2.3 Predictive model assessment

One of the major goals of statistical analysis is to make predictions along with reasonable measures of the uncertainty associated with them. Therefore, predictions should be probabilistic in nature, represented as probability distributions across future quantities and occurrences [34]. The goal of probabilistic forecasting is to maximize the sharpness of the predictive distributions subject to calibration [35]. Here, sharpness refers to the concentration of the predictive distributions, whereas calibration refers to the statistical consistency between the distributional forecasts and the observations [36].

One of the tools for predictive model assessment is the probability integral transform (PIT), which is simply the value that the predicted CDF attains at the observation. It is suggested that this tool for calibration checks for continuous data [34]. The PIT has a standard uniform distribution if the observation is taken from the predicted distribution, which is an ideal and desirable scenario. But the predictive distribution is discrete when it comes to count data and under the assumption of an ideal scenario, the PIT is no longer uniform. As a remedy, a randomized PIT was recommended [36], which follows:

$$u = P_{x-1} + v(P_x - P_{x-1}), \qquad (2.13)$$

where $P$ is the predictive distribution, $x\tilde{}P$ is the observed count, $v$ is standard uniform and independent of $x$ (Note: $P_{(-1)} = 0$ is standard uniform). Further, a nonrandomized uniform PIT has been proposed in which they replace the randomized PIT value in Eq. (2.13) by its conditional CDF given $x$, i.e.,

$$F(u|x) = \begin{cases} 0 & u \le P_{x-1}, \\ \frac{u-P_{x-1}}{P_x-P_{x-1}} & P_{x-1} \le u \ge P_x, \\ 1 & P_x \ge u. \end{cases} \qquad (2.14)$$

By comparing the mean PIT across a relevance set of $n$ predictions, calibration can be evaluated,

$$\bar{F}(u) = \frac{1}{n}\sum_{i=0}^{n} F^{(i)}(u|x^{(i)}), \quad 0 \le u \le 1.$$

They suggested to perform the comparison by nonrandomized PIT with height $f_j$, defined as:

$$f_j = \bar{F}(\frac{j}{J}) - \bar{F}(\frac{j-1}{J}), \;\; j = 1, 2, ..., J,$$

where $J$ denotes the number of bins of the histogram, which has been typically suggested as $J = 10$ or $J = 20$. So the calibration can be easily assessed graphically rather than numerically. That is, skewed histograms reflect biased central tendencies, whereas uniform, inversely U-shaped, and U-shaped histograms exhibit predicted distributions that are well-calibrated, over-dispersed, and under-dispersed, respectively.

## 3. Results

In this section, descriptive statistics of the included climate covariates as well as a graphical tool for predictive model assessment have been demonstrated to understand the nature of the raw data along with choosing an appropriate model. The overview of the summary information, such as minimum, maximum, mean, etc. for the daily dengue-infected patient and different climate variables across 1092 days is shown in Table 1.

**Table 1.** Summary statistics of dependent variable "daily dengue infected cases" and other climatic variables from 2021 to 2023.

| Variable | Mean (SD) | Minimum | Maximum |
|---|---|---|---|
| Dengue Infected Cases | 507.7 (2941.88) | 0 | 31461 |
| Min. temperature (°C) | 23.12 (4.63) | 10 | 30.7 |
| Max. temperature (°C) | 31.9 (3.82) | 15.4 | 40.6 |
| Wind speed | 2.45 (1.09) | 0 | 15 |
| Rainfall | 4.94 (15.25) | 0 | 255 |
| Visibility | 4.30 (0.71) | 0.4 | 6 |
| Humidity | 75.34 (14.36) | 27 | 100 |

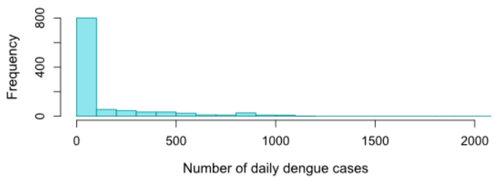


**Fig. 1.** Histogram of the daily dengue infected cases in Dhaka city from 2021 to 2023.

During the course of the study, there were, on average, 507.7 dengue-infected cases every day. Additionally, the minimum temperature of 23.12°C (with a maximum of 30.70°C and a minimum of 10°C), maximum temperature of 31.9 (with a maximum of 40.6°C and a minimum of 15.40°C), wind speed of 2.45 knots (with a maximum of 15 knots and a minimum of 0 knots), rainfall of 4.94 mm (with a maximum of 255 mm and a minimum of 0 mm), visibility of 4.30 km (with a maximum of 6 km and a minimum of 0.4 km), and humidity of 75.34% (with a maximum of 100% and a minimum of 27%), on average, have been observed from Table 1. Summary information also shows that the variance of the dengue-infected cases is 33.57 times the mean, which is the case of overdispersion. Moreover, the daily dengue-infected cases exhibit an overabundance of zero cases, as evident by the histogram in Fig. 1. Hence, overdispersion as well as zero inflation needs to be taken into account while modeling the daily dengue incidence. As a result, all possible count models, such as Poisson, negative binomial, discrete Lindley, discrete Weibull, zero-inflated Poisson, zero-inflated negative binomial, hurdle Poisson, hurdle negative binomial, etc. regression models were employed considered to analyze over-dispersed and zero-inflated dengue infected cases in our study. Based on various performance criteria, as listed in Table 2, the appropriate model has been selected.

One of the fundamental requirements for choosing a count model is dispersion, which, in the case of over-dispersed count response modeling, should have a desirable value of close to 1 for any count model. A dispersion value of near 1 indicates that the particular model adequately captures overdispersion. From the results of Table 2, it is seen that dispersion for negative binomial, discrete Weibull, zero-inflated negative binomial, and hurdle negative binomial are 2.97, 0.99, 3.08, 1.91, respectively, whereas for Poisson, discrete Lindley, zero-inflated Poisson, and hurdle Poisson, the values are 7469.92, 24.65, 80.92, and 79.87, respectively. So, an obvious indication of the dispersion value is that discrete Weibull incorporates overdispersion quite well in comparison to the other count models.

Further, average AIC (AIC) values have also been considered as a performance criterion. The lower the AIC, the better the model. As evidenced in Table 2, lower AIC values have been found for negative binomial (AIC: 9.68), discrete Weibull (AIC: 9.55), zero-inflated negative binomial (AIC: 9.67), and hurdle negative binomial (AIC: 9.45) regression models as compared to the other models in this study, which is consistent with the conclusion of dispersion values.

Moreover, Chi-square test for good-

**Table 2.** Goodness-of-fit tests based on AIC, Pearson residual $\chi^2$, $p$-value, and Dispersion for fitted count regression models.

| Model | AIC | DF | Pearson $\chi^2$ | $p$-value | Dispersion |
|---|---|---|---|---|---|
| PR | 2200.14 | 1085 | 8104865 | <0.001 | 7469.92 |
| NBR | 9.68 | 1085 | 3221.608 | <0.001 | 2.97 |
| DL | 15.17 | 1085 | 26743.44 | <0.001 | 24.65 |
| DW | **9.55** | 1084 | 1042.393 | **0.805** | **0.99** |
| ZIP | 2031.01 | 1078 | 87236.22 | <0.001 | 80.92 |
| ZINB | 9.67 | 1077 | 3317.111 | <0.001 | 3.08 |
| HP | 2031.02 | 1078 | 86097.46 | <0.001 | 79.87 |
| HNB | 9.45 | 1077 | 2052.521 | <0.001 | 1.91 |

**AIC: average AIC**

ness of fit, the $H_0$: Observed data comes from discrete Weibull distribution versus $H_1$: the data does not come from discrete Weibull distribution. The larger p-value of the Pearson residual Chi-square indicates that we fail to reject the null hypothesis, implying that observed data may come from discrete Weibull distribution. Therefore, our results suggest that the discrete Weibull (DW) model provides an adequate and reliable fit to the dengue incidence data. In addition to the goodness of fit, PIT, a graphical tool has been utilized to help determine the predictive model with greater accuracy. The more deviation from uniformity, the less inadequate the model is. A histogram close to uniform (0,1) refers to a well-calibrated predictive model, whereas a biased central tendency, either under or over-dispersed, is indicated by a skewed, U-shaped, or inverse U-shaped PIT histogram in the predicted model. From Fig. 2, it is clearly visible that the PIT histogram is also consistent with the results of AIC. For instance, the PIT histograms of the models having lower AIC Values, including Negative Binomial, discrete Weibull, zero-inflated negative binomial and hurdle negative binomial, are roughly close to uniformity, as evident by the histogram in Fig. 2. This indicates that these models are approximately calibrated when they are fitted to actual datasets. But except discrete Weibull, the other models don't capture dispersion as well as zero-inflation ratio properly. Thus, combining results from both model selection procedures, the discrete Weibull model has been considered the appropriate model over all other models in this study. Therefore, in order to identify the significant climate factors influencing dengue incidence, only regression outputs obtained from the discrete Weibull regression model have been taken into consideration, with a $p$-value of less than 0.05 serving as a benchmark.

As can be seen from Table 3, climatic factors including min. temperature, max. temperature, wind speed, and humidity are found to have significant effects with very low $p$-value (<0.001) on daily dengue infected cases, whereas rainfall and visibility do not have to actual datasets. But except discrete Weibull, the other models don't capture dispersion as well as zero-inflation ratio properly.

Thus, combining results from both model selection procedures, the discrete Weibull model has been considered the appropriate model over all other models in this study. Therefore, in order to identify the significant climate factors influencing

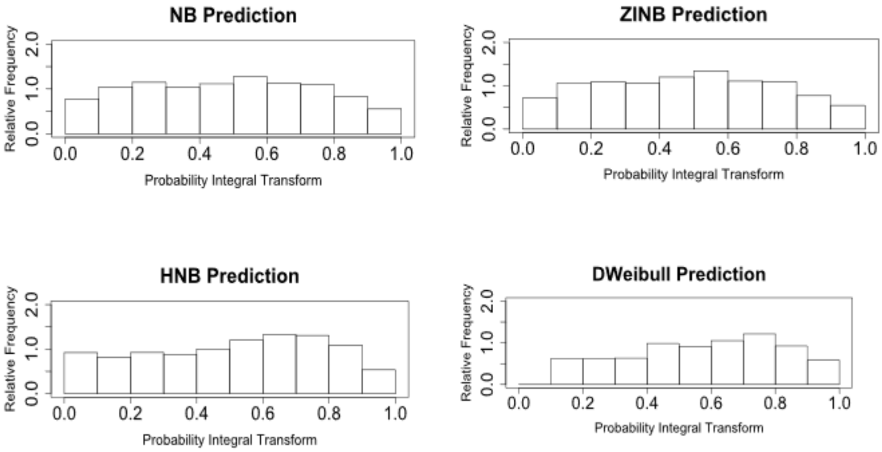


**Fig. 2.** Probability Integral Transform (PIT) histograms.

**Table 3.** Outputs for discrete Weibull regression model for daily dengue infected cases over the period of 3 years, 2021 to 2023.

| Variable | Estimate | SE | IRR | 95% CI for IRR | *p*-value |
|---|---|---|---|---|---|
| Intercept | -6.22 | 0.925 | 0.002 | (-0.002, 0.005) | <0.001*** |
| Min. temperature | 0.25 | 0.03 | 1.284 | (1.21, 1.36) | <0.001*** |
| Max. temperature | -0.121 | 0.048 | 0.886 | (0.80, 0.97) | 0.012* |
| Wind speed | -0.346 | 0.066 | 0.707 | (0.62, 0.79) | <0.001*** |
| Rainfall | -0.004 | 0.004 | 0.996 | (0.99, 1.01) | 0.357 |
| Visibility | 0.304 | 0.212 | 1.355 | (0.79, 1.92) | 0.153 |
| Humidity | 0.098 | 0.006 | 1.103 | (1.09, 1.12) | <0.001*** |

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

dengue incidence, only regression outputs obtained from the discrete Weibull regression model have been taken into consideration, with a *p*-value of less than 0.05 serving as a benchmark.

As can be seen from Table 3, climatic factors including min. temperature, max. temperature, wind speed, and humidity are found to have significant effects with very low *p*-value (<0.001) on daily dengue infected cases, whereas rainfall and visibility do not have statistically significant effects on dengue incidence. Results from the incidence risk ratio (IRR) presented in Table 3 show that when all predictors (min. temperature, max. temperature, wind speed, rainfall, visibility, and humidity) are equal to zero, the rate of dengue-infected cases decreases by 99.8%. Further, keeping all other climatic factors at a fixed level, an increase in min. temperature of just 1°C increases the expected dengue incidence by 28.4%, while an increase in max. temperature of just 1°C decreases the expected incidence rate by 11.4%. Similar to max. temperature, wind speed has a negative impact on dengue-infected cases. For every unit increase in wind speed, the expected number of dengue-infected cases decreases by 29.3%, assuming the values of the other relevant predictor variables remain constant. Conversely, a 1% rise in humidity causes

**Table 4.** Goodness-of-fit tests for the fitted count regression models on simulated data.

| Model | AIC | DF | Pearson $\chi^2$ | $p$-value | Dispersion |
|---|---|---|---|---|---|
| NB | 9.65 | 1085 | 2214.76 | <0.001 | 2.04 |
| DW | 9.57 | 1084 | 1067.59 | 0.633 | 0.98 |
| ZINB | 9.67 | 1077 | 2198.09 | <0.001 | 2.04 |
| HNB | 9.62 | 1077 | 1745.51 | <0.001 | 1.62 |

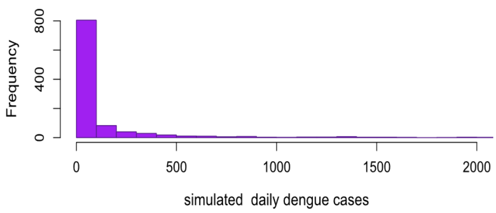


**Fig. 3.** Histogram of the simulated dengue cases.

the expected incidence rate of dengue to increase to 1.103 times holding other significant covariates as fixed.

## 4. Simulation Study

To corroborate the findings obtained from real-life application on daily dengue cases, a simulation study was conducted. Given that the discrete Weibull was identified as the most suitable among the fitted models based on the characteristics of the dataset, we configured a simulation setting to assess the robustness of this conclusion. The primary objective of this study was to simulate daily dengue counts from the discrete Weibull distribution that closely mimic the real dataset and validate the performance of the models using real life application.

Therefore, the DW regression model was fitted on real dengue dataset using the following link function $\log(-\log(q_i)) = \beta_0 + \beta_1 \times min.temperature + \beta_2 \times max.temperature + \beta_3 \times windspeed + \beta_4 \times rainfall + \beta_5 \times visibility + \beta_6 \times humidity$. Hence, the estimated regression parameters: $\beta = (\beta_0, \beta_1, \beta_2, \beta_3, \beta_4, \beta_5, \beta_6)^T$ = (-6.22, 0.25,-0.12, -0.35, -0.004, 0.30, 0.098) with shape parameter, $\alpha = 0.403$, which have been further used to generate the count data of the same sample size as the original one. Subsequently, to verify the data structure, the histogram of the generated count responses (Fig. 3) was evaluated, which exhibited patterns of over-dispersion consistent with those observed in the real dataset. Afterwards, the NB, ZINB, HNB, and DW regression models were fitted on the simulated data, and DW was found as the most suitable model, on the basis of the values of AIC, $p$-value of Pearson residual Chi-square, and dispersion (Table 4). DW was also affirmed the most suitable model among the fitted models for the simulated data by the predictive model assessment tool (Fig. 4).

## 5. Discussion

Dengue fever has been a major public health threat across many countries of the world, including Bangladesh. There have been many analyses on dengue transmission through several meteorological as well as demographic variables. This study aimed at focusing on influential determinants of dengue infected cases in Dhaka city, Bangladesh over the study period of 3 years, from 2021 to 2023. Due to the excess variation across responses as well as the over-abundance of zero counts in the dengue infected cases, almost all possible count regression models have been employed, among which discrete Weibull regression model has been chosen as the best fitted model for our study based on the values of AIC, dispersion and the graphical approach.

The findings of this study suggest that particular environmental factors, such as minimum temperature, maximum temperature, wind speed, and humidity are statistically significant in determining the

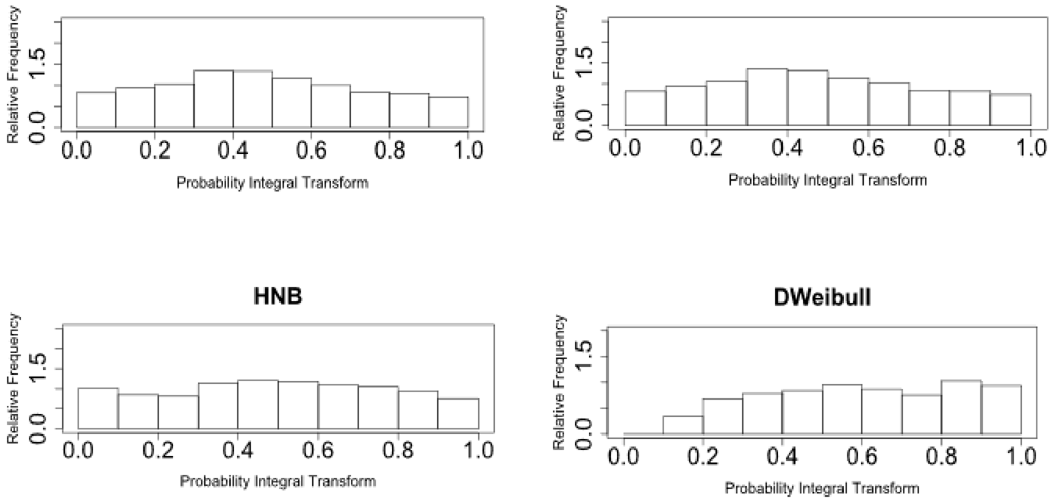


**Fig. 4.** Probability Integral Transform (PIT) histograms on simulated data.

spread of dengue. On the other hand, rainfall and visibility had no effect on dengue occurrences. Aedes mosquito growth usually occurs best at temperatures between 25°C and 30°C [12]. The fast growth and development of mosquito larvae at high temperatures causes the adult mosquitoes to emerge earlier. While the development process may slow down in lower temperatures, it may accelerate in warmer ones. This study also confirmed this relationship of temperature and dengue incidence in two opposite directions. The min. temperature range of 10°C to 30.70°C observed during the study period had a positive association with dengue incidence. This finding is consistent with the results of other studies conducted in Bangladesh [5], India [38], Taiwan [25], Malaysia [20], Northern Italy [39], and Indonesia [40], where the authors also reported a positive relationship between temperature and dengue outbreaks. In contrast, the max. temperature showed an opposite result, which is in line with some previous studies [22, 23, 26].

Excessively high wind speeds prevent adult mosquitoes from flying, which reduces their propensity to bite [12], which is supported by our study findings. Wind speed being a statistically significant environmental factor for dengue cases, decreases the spread of dengue incidence. From our study, average wind speed of 2.45 knots has been observed and dengue incidence rate falls by 29.3% for 1 knot increase in wind speed. The highest incidence of dengue fever was observed with wind speeds of 5-6 knots [44]. Regarding another study, there is inconsistency as wind speed had a positive impact on dengue incidence [5].

According to our study findings, rainfall has no significant effect on dengue outbreaks having consistency with several studies [13, 19, 24] but studies in Indonesia [22], French Guiana [41], and Thailand [42] showed a negative effect of rainfall on dengue incidence. This might be because excessive rainfall can wash away mosquito larvae from breeding sites, leading to a reduction in mosquito populations. Dengue incidence may be indirectly influenced by visibility, which is defined as atmospheric clarity or the distance at which things are explicitly seen. When visibility is low, people could be less inclined to go outside or engage in outdoor activities, which would limit their exposure to mosquito bites. But results from this study indicate visibility as an insignificant risk factor for dengue transmission, which could be that Aedes mosquitoes, the primary dengue vector, are day-biters and their breeding is less affected by visibility changes compared to nocturnal mosquitoes.

An abundance of humidity creates ideal conditions for Aedes mosquito reproduction. High humidity can extend the life of the eggs and larvae laid by these mosquitoes in stagnant water, hastening the development of the larvae into adults. As a result, humid environments are better for mosquito reproduction, raising the possibility of dengue spreading, which also validates our outcomes. Several studies also have similar results [5, 13, 43, 44]. Contradictory findings were also observed [22, 45].

## 6. Conclusion

This study investigates the association between climate factors and dengue incidence in Dhaka, Bangladesh. Our research affirms a significant correlation between dengue transmission and environmental factors, specifically minimum temperature, maximum temperature, wind speed, and humidity. Minimum temperature and humidity demonstrate a positive influence, whereas wind speed shows a negative impact. Rainfall and visibility, however, do not appear to affect dengue transmission according to our study. These findings can guide policymakers in establishing a climate-based warning system. Subsequent endeavors may involve constructing a comprehensive model integrating immunological, entomological, demographic, and climatic data to predict dengue cases accurately.

Our study is subject to certain limitations. While we have focused solely on climatic factors, it's important to acknowledge that demographic and socio-economic factors may also play a role in shaping the prevalence of dengue incidence. Additionally, beyond the six climatic factors we have analyzed, there are other climate variables, such as sunshine hours and seasonal variability, which could further enrich our investigation if data on these variables could be collected. Integrating these factors into our analysis could provide a more comprehensive understanding of the dynamics influencing dengue transmission. Additionally, under-reported dengue cases can also be considered. Moreover, non-linearity, machine learning techniques along with spatial analysis could be incorporated.

## References

[1] Bhatt S, Gething PW, Brady OJ, Messina JP, Farlow AW, Moyes CL, Drake JM, Brownstein JS, Hoen AG, Sankoh O, Myers MF. The global distribution and burden of dengue. Nature. 2013 Apr 25;496(7446):504-7.

[2] Stanaway JD, Shepard DS, Undurraga EA, Halasa YA, Coffeng LE, Brady OJ, Hay SI, Bedi N, Bensenor IM, Castañeda-Orjuela CA, [[23]] Chuang TW. The global burden of dengue: an analysis from the Global Burden of Disease Study 2013. The Lancet infectious diseases. 2016 Jun 1;16(6):712-23.

[3] World Health Organization. Dengue and severe dengue [Internet]. Geneva: [[3]] World Health Organization; 2020 [cited 2024 Oct 2].

[4] Karim MN, Munshi SU, Anwar N, Alam MS. Climatic factors influencing dengue cases in Dhaka city: a model for dengue

prediction. Indian journal of medical research. 2012 Jul 1;136(1):32-9.

[5] Hossain S, Islam MM, Hasan MA, Chowdhury PB, Easty IA, Tusar MK, Rashid MB, Bashar K. Association of climate factors with dengue incidence in Bangladesh, Dhaka City: a count regression approach. Heliyon. 2023 May 1;9(5).

[6] Amin MM, Hussain AM, Nahar K, Chowdhury IA, Murshed M, Chowdhury SA. Searo-diagnosis of Dengue Infections in Four Metropolitan Cities of Bangladesh.

[7] Sharmin S, Glass K, Viennet E, Harley D. Interaction of mean temperature and daily fluctuation influences dengue incidence in Dhaka, Bangladesh. PLoS neglected tropical diseases. 2015 Jul 10;9(7):e0003901.

[8] Abir T, Ekwudu OM, Kalimullah NA, Nur-A Yazdani DM, Al Mamun A, Basak P, Osuagwu UL, Permarupan PY, Milton AH, Talukder SH, Agho KE. Dengue in Dhaka, Bangladesh: hospital-based cross-sectional KAP assessment at Dhaka north and Dhaka south city corporation area. PLoS One. 2021 Mar 30;16(3):e0249135.

[9] Bonna AS, Pavel SR, Mehjabin T, Ali M. Dengue in bangladesh. IJID One Health. 2023 Dec 1;1:100001.

[10] Akram A. Alarming turn of dengue fever in Dhaka city in 2019. Bangladesh Journal of Infectious Diseases. 2019;6(1):1.

[11] Sajib AH, Akter S, Saha G, Hossain Z. Demographic-environmental effect on dengue outbreaks in 11 countries. Plos one. 2024 Sep 11;19(9):e0305854.

[12] Karim MR, Sarker I, Bashar K. Changing patterns of climatic risk factors for the transmission of dengue fever in Bangladesh: count model approach. Journal of Public Health. 2023 Nov 13:1-0.

[13] Banu S, Guo Y, Hu W, Dale P, Mackenzie JS, Mengersen K, Tong S. Impacts of El Niño southern oscillation and Indian Ocean dipole on dengue incidence in Bangladesh. Scientific reports. 2015 Nov 5;5(1):16105.

[14] Meem SM, Hossain MT, Chowdhury JK, Miah MS, Monir MF. Understanding the dynamics of dengue in Bangladesh: EDA, climate correlation, and predictive modeling. InTENCON 2023-2023 IEEE Region 10 Conference (TENCON) 2023 Oct 31 (pp. 1309-1314). IEEE.

[15] Sharmin S, Viennet E, Glass K, Harley D. The emergence of dengue in Bangladesh: epidemiology, challenges and future disease risk. Transactions of The Royal Society of Tropical Medicine and Hygiene. 2015 Oct 1;109(10):619-27.

[16] Bal S, Sodoudi S. Modeling and prediction of dengue occurrences in Kolkata, India, based on climate factors. International journal of biometeorology. 2020 Aug;64:1379-91.

[17] Malik A, Yasar A, Tabinda AB, Zaheer IE, Malik K, Batool A, Mahfooz Y. Assessing spatio-temporal trend of vector breeding and dengue fever incidence in association with meteorological conditions. Environmental monitoring and assessment. 2017 Apr;189:1-20.

[18] Shabbir W, Pilz J, Naeem A. A spatial-temporal study for the spread of dengue depending on climate factors in Pakistan (2006–2017). BMC Public Health. 2020 Dec;20:1-0.

[19] Tuladhar R, Singh A, Varma A, Choudhary DK. Climatic factors influencing dengue incidence in an epidemic area of Nepal. BMC research notes. 2019 Dec;12:1-7.

[20] Cheong YL, Burkart K, Leitão PJ, Lakes T. Assessing weather effects on dengue

disease in Malaysia. International journal of environmental research and public health. 2013 Dec;10(12):6319-34.

[21] Xuan LT, Van Hau P, Thu DT, Toan DT. Estimates of meteorological variability in association with dengue cases in a coastal city in northern Vietnam: an ecological study. Global health action. 2014 Dec 1;7(1):23119.

[22] Tosepu R, Tantrakarnapa K, Nakhapakorn K, Worakhunpiset S. Climate variability and dengue hemorrhagic fever in Southeast Sulawesi Province, Indonesia. Environmental Science and Pollution Research. 2018 May;25:14944-52.

[23] Iguchi JA, Seposo XT, Honda Y. Meteorological factors affecting dengue incidence in Davao, Philippines. BMC public health. 2018 Dec;18:1-0.

[24] Ostovar A, Haghdoost AA, Rahimiforoushani A, Raeisi A, Majdzadeh R. Time series analysis of meteorological factors influencing malaria in South Eastern Iran. Journal of arthropod-borne diseases. 2016 Jun;10(2):222.

[25] Chuang TW, Chaves LF, Chen PJ. Effects of local and regional climatic fluctuations on dengue outbreaks in southern Taiwan. PLoS One. 2017 Jun 2;12(6):e0178698.

[26] Duarte JL, Diaz-Quijano FA, Batista AC, Giatti LL. Climatic variables associated with dengue incidence in a city of the Western Brazilian Amazon region. Revista da Sociedade Brasileira de Medicina Tropical. 2019 Feb 21;52:e20180429.

[27] Akter S, Sharmin S, Akter T, Sajib AH. Comparing the Performance of Count Regression Models to Assess the Impact of Climate on COVID-19 Incidence in Dhaka, Bangladesh. International Journal of Statistical Sciences. 2024 Dec 23:149-62.

[28] Hilbe JM. Negative binomial regression. Cambridge University Press; 2011.

[29] Maxwell O, Mayowa BA, Chinedu IU, Peace AE. Modelling count data; a generalized linear model framework. Am J Math Stat. 2018;8(6):179-83.

[30] Gómez-Déniz E, Calderín-Ojeda E. The discrete Lindley distribution: properties and applications. Journal of statistical computation and simulation. 2011 Nov 1;81(11):1405-16.

[31] Nguyen MV, Le NT. Using the Discrete Lindley Distribution to Deal with Overdispersion in Count Data. Austrian Journal of Statistics. 2023 Jul 18;52(3):96-113.

[32] Klakattawi HS, Vinciotti V, Yu K. A simple and adaptive dispersion regression model for count data. Entropy. 2018 Feb 22;20(2):142.

[33] Mullahy J. Specification and testing of some modified count data models. Journal of econometrics. 1986 Dec 1;33(3):341-65.

[34] Dawid AP. Present position and potential developments: Some personal views statistical theory the prequential approach. Journal of the Royal Statistical Society: Series A (General). 1984 Mar;147(2):278-90.

[35] Gneiting T, Balabdaoui F, Raftery AE. Probabilistic forecasts, calibration and sharpness. Journal of the Royal Statistical Society Series B: Statistical Methodology. 2007 Apr;69(2):243-68.

[36] Czado C, Gneiting T, Held L. Predictive model assessment for count data. Biometrics. 2009 Dec;65(4):1254-61.

[37] Smith JQ. Diagnostic checks of non□standard time series models. Journal of Forecasting. 1985;4(3):283-91.

[38] Mutheneni SR, Morse AP, Caminade C, Upadhyayula SM. Dengue burden in India: recent trends and importance of cli-

matic parameters. Emerging microbes & infections. 2017 Jan 1;6(1):1-0.

[39] Roiz D, Rosa R, Arnoldi D, Rizzoli A. Effects of temperature and rainfall on the activity and dynamics of host-seeking Aedes albopictus females in northern Italy. Vector-borne and zoonotic diseases. 2010 Oct 1;10(8):811-6.

[40] Arcari P, Tapper N, Pfueller S. Regional variability in relationships between climate and dengue/DHF in Indonesia. Singapore Journal of Tropical Geography. 2007 Nov;28(3):251-72.

[41] Adde A, Roucou P, Mangeas M, Ardillon V, Desenclos JC, Rousset D, Girod R, Briolant S, Quenel P, Flamand C. Predicting dengue fever outbreaks in French Guiana using climate indicators. PLoS neglected tropical diseases. 2016 Apr 29;10(4):e0004681.

[42] Thammapalo S, Chongsuwiwatwong V, McNeil D, Geater A. The climatic factors influencing the occurrence of dengue hemorrhagic fever in Thailand. Southeast Asian J Trop Med Public Health. 2005 Jan 1;36(1):191-6.

[43] Sirisena PD, Noordeen F, Kurukulasuriya H, Romesh TA, Fernando L. Effect of climatic factors and population density on the distribution of dengue in Sri Lanka: a GIS based evaluation for prediction of outbreaks. PloS one. 2017 Jan 9;12(1):e0166806.

[44] Baharom M, Ahmad N, Hod R, Arsad FS, Tangang F. The impact of meteorological factors on communicable disease incidence and its projection: a systematic review. International Journal of Environmental Research and Public Health. 2021 Oct 22;18(21):11117.

[45] Sehgal M, Ghosh S. Exploring the Usefulness of Meteorological Data for Predicting Malaria Cases in Visakhapatnam, Andhra Pradesh. Weather, Climate, and Society. 2020 Apr 1;12(2):323-30.