Title

Prediction of dengue annual incidence using seasonal climate variability in Bangladesh between 2000 and 2018.

Journal

PLOS Global Public Health, 2(5)

Authors

Hossain, M
Zhou, Wen
Ren, Chao
et al.

Publication Date

2022

DOI

10.1371/journal.pgph.0000047

Peer reviewed

PLOS GLOBAL PUBLIC HEALTH

RESEARCH ARTICLE

Prediction of dengue annual incidence using seasonal climate variability in Bangladesh between 2000 and 2018

M. Pear Hossain[1,2], Wen Zhou[3], Chao Ren[4], John Marshall[5], Hsiang-Yu Yuan[1]*

1 Department of Biomedical Sciences, Jockey Club College of Veterinary Medicine and Life Sciences, City University of Hong Kong, Kowloon, Hong Kong, **2** Department of Statistics, Bangabandhu Sheikh Mujibur Rahman Science and Technology University, Gopalganj, Bangladesh, **3** School of Energy and Environment, City University of Hong Kong, Kowloon, Hong Kong, **4** Faculty of Architecture, The University of Hong Kong, Pokfulam, Hong Kong, **5** Division of Biostatistics, School of Public Health, University of California, Berkeley, California, United States of America

* sean.yuan@cityu.edu.hk

Citation: Hossain MP, Zhou W, Ren C, Marshall J, Yuan H-Y (2022) Prediction of dengue annual incidence using seasonal climate variability in Bangladesh between 2000 and 2018. PLOS Glob Public Health 2(5): e0000047. https://doi.org/10.1371/journal.pgph.0000047

Editor: Raph L. Hamers, University of Oxford, INDONESIA

Received: May 28, 2021

Accepted: October 18, 2021

Published: May 9, 2022

Data Availability Statement: All data are in the manuscript and/or Supporting information files.

Funding: Grant number: #7200573 and #9610416 Name: City University of Hong Kong URL: https://www.cityu.edu.hk/ Recipient: Hsiang-Yu Yuan The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: The authors have declared that no competing interests exist.

Abstract

The incidence of dengue has increased rapidly in Bangladesh since 2010 with an outbreak in 2018 reaching a historically high number of cases, 10,148. A better understanding of the effects of climate variability before dengue season on the increasing incidence of dengue in Bangladesh can enable early warning of future outbreaks. We developed a generalized linear model to predict the number of annual dengue cases based on monthly minimum temperature, rainfall and sunshine prior to dengue season. Variable selection and leave-one-out cross-validation were performed to identify the best prediction model and to evaluate the model's performance. Our model successfully predicted the largest outbreak in 2018, with 10,077 cases (95% CI: [9,912–10,276]), in addition to smaller outbreaks in five different years (2003, 2006, 2010, 2012 and 2014) and successfully identified the increasing trend in cases between 2010 and 2018. We found that temperature was positively associated with the annual incidence during the late winter months (between January and March) but negatively associated during the early summer (between April and June). Our results might be suggest an optimal minimum temperature for mosquito growth of 21–23˚*C*. This study has implications for understanding how climate variability has affected recent dengue expansion in neighbours of Bangladesh (such as northern India and Southeast Asia).

Background

Dengue fever, one of the most prevalent vector-borne diseases, has led to significant socio-economic costs in many parts of the world [1]. Three-quarters of the global dengue cases occur in Southeast Asian and western Pacific countries, due to the associated favorable weather conditions for mosquito population expansion [2, 3]. Outbreaks of dengue fever can significantly reduce life expectancy due to the possibility of developing severe dengue following secondary infections from different dengue serotypes [4]. Therefore, it is critically important to understand the impacts of the climate on the spread of dengue in these regions, as this can serve as early warning system and enable early preventative measures to be put in place before outbreaks become established.

Expansion of dengue in the regions surrounding northern India may have occurred in recent years due to climate change. The prolonged rainy seasons and increasing temperatures in subtropical regions of Southeast Asia may provide favorable conditions for expansion of *Aedes* mosquito populations, the dengue vector [5–8]. In addition, increased incidence of dengue has recently been observed in more temperate regions [9], such as Nepal [10], indicating a possible expansion of the disease from the subtropics to cooler climates, posing a threat to northern India, Pakistan and their neighbors. Bangladesh is located to the northeast of India and to the south of Nepal, and lies along the Tropic of Cancer. Understanding the patterns of recent dengue outbreaks in Bangladesh may provide greater insight into whether dengue has expanded into the region surrounding northern India, a region with more than 140 million inhabitants.

Dengue fever was first identified in Bangladesh in 1964 [11] and was not initially considered to be a severe threat to public health. However, in 2000, an outbreak occurred, leading to a total of 5,551 reported cases and 93 confirmed deaths [12, 13]. The average annual number of dengue cases decreased between 2000 and 2010. However, since then the number of annual dengue cases in Bangladesh has been increasing rapidly. A recent outbreak in 2019 was the largest ever experienced by the country, whereas the second largest outbreak was seen only a year prior, in 2018. Whether or how climate variability may have driven this unprecedented rise in outbreak size in 2018 and 2019 is still largely unknown [14].

Several studies have been carried out to estimate dengue incidence in Bangladesh using climate data prior to 2010 [15–17]; however, the driving factors responsible for the increasing disease burden since 2010 remain to be investigated. In these previous studies, temperature and rainfall were found to be significant contributing factors [18–21]. Previous studies also assumed that the effects of climate variables are independent of the time of year. However, the effects of climate variables can also be time-dependent. Several studies have demonstrated that the effects of rainfall on dengue incidence can vary throughout the year [22, 23]. The abundant rainfall that occurs during monsoon season is likely to have negative effects on mosquito population size, as the rain can disrupt potential mosquito habitats. In contrast, rainfall in winter months may result in stagnant bodies of water suitable for mosquito breeding.

Recent studies have mentioned the dengue incidence and mosquito abundance can be affected by weather conditions up to 5 months before the season starts [22–24] using data in or near subtropical areas. However, most of the studies focus on climate factors during dengue season. A better understanding of the relationship between climate variability before dengue season and annual incidence can provide insight into whether dengue has been expanding in a region and allow early warning system to be built.

This study aimed to estimate the effects of climate factors before dengue season on annual incidence in Bangladesh using historical data from 2000 to 2018. We developed a generalized linear model to predict annual dengue cases based on monthly temperature, rainfall and sunshine. We demonstrated that temperature and rainfall have variable effects on dengue incidence depending on the time of year and suggest an ideal temperature range for mosquito population growth based on our findings.

Methods

Study location

Bangladesh is a Southeast Asian country, as defined by the World Health Organization (WHO) [25]. India surrounds it to the east, west and north, and Myanmar borders it to the

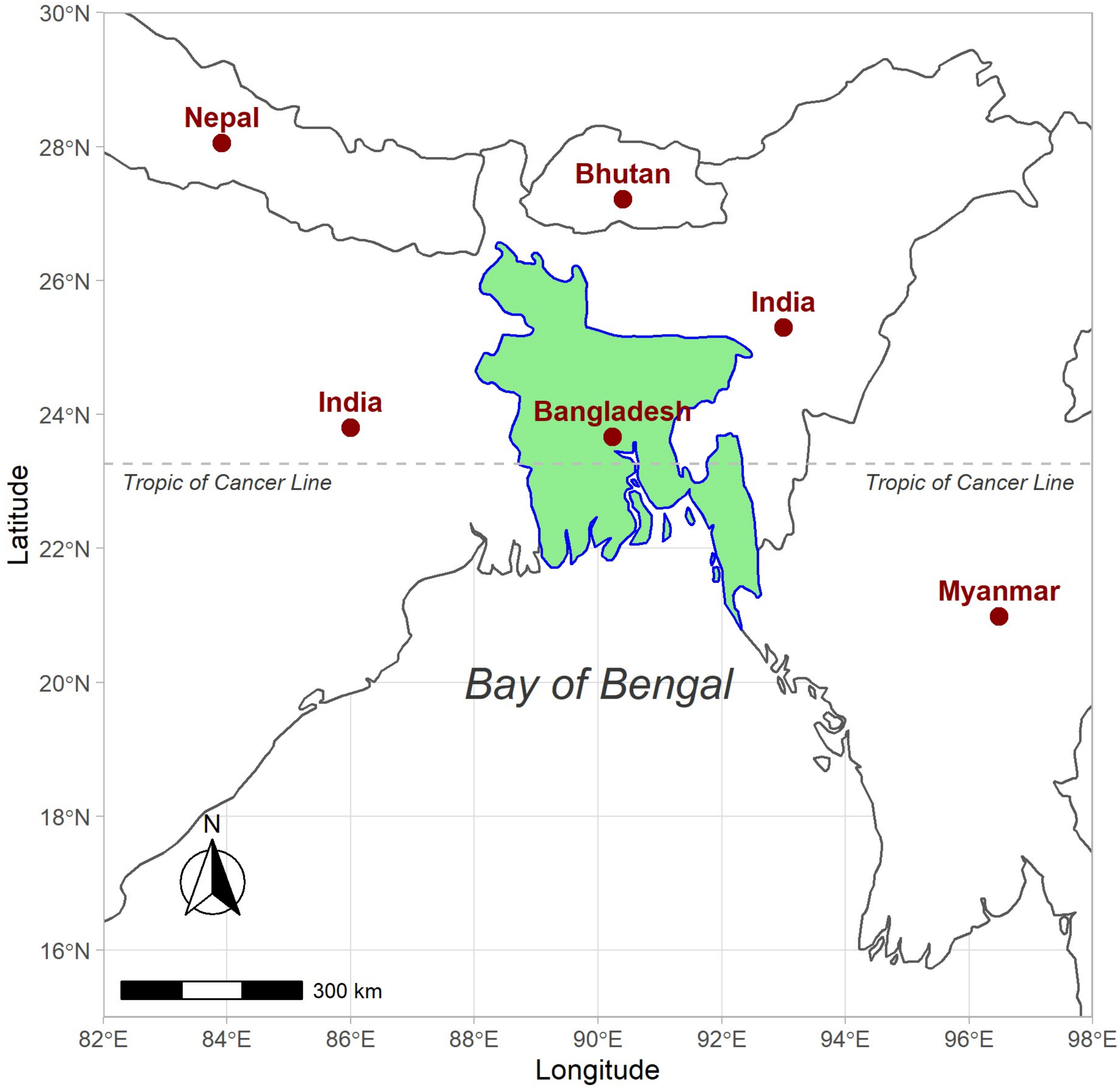

Fig 1. Location of Bangladesh on the world map. India surrounds Bangladesh on three sides (east, west and north) and the Bay of Bengal is to the south. The country shares a border with Myanmar in the southeast. The Tropic of Cancer line crosses the middle of the country.

https://doi.org/10.1371/journal.pgph.0000047.g001

southeast (Fig 1). The Bay of Bengal is located in the south. Bangladesh is located at $20^{\circ}59'N$ to $26^{\circ}63'N$ and $88^{\circ}03'E$ to $92^{\circ}67'E$. The Tropic of Cancer line is located at $23^{\circ}26'N$ and $88^{\circ}47'E$, where it crosses Bangladesh from east to west [26].

Bangladesh is located in both tropical and subtropical climate regions. The seasons in Bangladesh can be broadly characterized as summer (March–June; mostly hot and humid), monsoon (June–October; warm and rainy) and winter (October–March; cold and dry). However, March can also be described as the spring, and the duration between mid-October and mid-November can be called the autumn. The maximum temperature ranges from $30^{\circ}C$ to $40^{\circ}C$ during summer, whereas in winter the average temperature reaches as low as $10^{\circ}C$ in most areas of the country. The average annual rainfall ranges between 1,500 mm and 3,000 mm.

Approximately 70–80% of the annual rainfall occurs during monsoon season [27]. The shortest period of sunshine, 5.4–5.8 hours per day, also occurs during this season. In contrast, winter and summer have the longest sunshine duration, 8.5–9.1 hours per day [28].

Dengue data

Dengue cases observed in health facilities across the country are generally reported to the Directorate General of Health Services (DGHS), and are classified into suspected, probable and confirmed cases. Individuals having acute febrile illness with or without non-specific signs and symptoms are classified as suspected cases and those having acute febrile illness with serological diagnosis are considered probable cases. The confirmed cases should have an acute febrile illness with positive dengue NS1 antigen or PCR test. Details of dengue case definitions and management are available from the DGHS [29].

The communicable disease control (CDC) unit of the DGHS compiles the reported dengue cases on a daily basis for further circulation. We accessed monthly dengue cases between January 2000 and December 2018 from the DGHS by collaborating with the Institute of Epidemiology, Disease Control and Research.

Climate data

Accumulated weather information was monitored and managed by the Bangladesh Meteorological Department (BMD) at 35 distinct weather stations across the country (location of the stations are available in Fig 1 in [15]). We collected these weather records from the BMD including daily mean, minimum and maximum temperature, total and maximum daily rainfall and daily sunshine duration. Temperature and rainfall are measured in degree Celsius ($^{\circ}C$) and millimeter (mm), respectively, whereas sunshine duration is recorded in hours. Daily information was averaged for each month to obtain monthly information for each station. The national averages for monthly temperature, sunshine duration and rainfall were obtained by averaging the values of all 35 weather stations.

Model formulation

Following the previous approaches, we modeled annual dengue incidence using quasi Poisson regression. In order to deal with the potential overdispersion issue, the quasi Poisson model and corrected quasi Akaike information criteria (QAICc) were used. QAICc has been frequently adopted for estimating the goodness of fit in modeling overdispersed count data in biological or ecological studies using the quasi Poisson regression [30]. Another approach is to adopt a negative binomial model. Hence, we have adopted QAICc in quasi Poisson regression model and also evaluate whether a negative binomial regression model can provide a best prediction result.

We assumed that dengue cases reported in January, February and March were belonging to the previous year's dengue outbreak. Hence, annual dengue incidence is defined as the sum of the number of dengue cases from April to December of a given year and from January to March of the following year. Let y_j be the annual dengue cases in the j^{th} year ($j = 1, 2, \ldots, 19$) such that $y_j \sim$ quasi- Poisson(λ_j), where λ_j represents the expected number of dengue cases in the j^{th} year, i.e. $E(y_j) = \lambda_j$. Therefore, the Poisson regression model can be expressed as

$$\log(\lambda_j) = \alpha + \sum_{i=1}^{6}\beta_i T_{ij} + \sum_{i=1}^{6}\gamma_i R_{ij} + \sum_{i=4}^{6}\eta_i S_{ij}, \tag{1}$$

where β_i, η_i and γ_i represent coefficients of temperature (T), rainfall (R) and sunshine duration (S) in the i^{th} month ($i = 1, 2, \ldots, 6$). Therefore, 15 predictor variables exist in the full model as shown in Eq (1). These predictors were preselected during these months as we aimed to predict dengue outbreaks, which often begin during the early summer. The final model was selected based on QAICc, which was designed to deal with small samples, and the results of cross-validation (details are given in Model selection and validation sections). It is worth to note that this approach has been previously applied in the annual dengue prediction in Asian countries [22, 23].

Model selection

We compared six different models for predicting the annual dengue outbreaks with various combinations of climate variables. Monthly temperatures (minimum, average or maximum) and rainfall (maximum or total) from January to June and the sunshine duration from April to June were chosen as potential predictors. Sunshine duration is mainly related to mosquito activities (e.g., mosquito bite) and hence the associated disease transmissibility. As the number of mosquitoes between January and March is low there is no need to consider sunshine duration during this period. Hence, we considered a shorter window for the sunshine duration. On the other hand, temperature and rainfall are involved in mosquito population growth. The change of growth rate during early months (e.g. from January to March) can affect population size later. To avoid overlapping the climate predictors, we considered a single category of temperatures and rainfall in each model.

A corrected version of the Akaike information criteria (AIC_c) [31] was used to extract potential predictor variables. The best prediction model was determined using a two-stage selection approach (Fig 2). In the first stage, variable selection was performed using forward stepwise AIC_c selection for each of the models. In each step of the stepwise selection, we recorded AIC_c along with parameter estimates, mean squared errors for validation and training to assess the fitness of the models (the details of the forward stepwise selection approach are given in the supplementary section). In the final step, each model provides a set of variables with minimum AIC_c. For these models, the model with the lowest AIC_c was selected as the

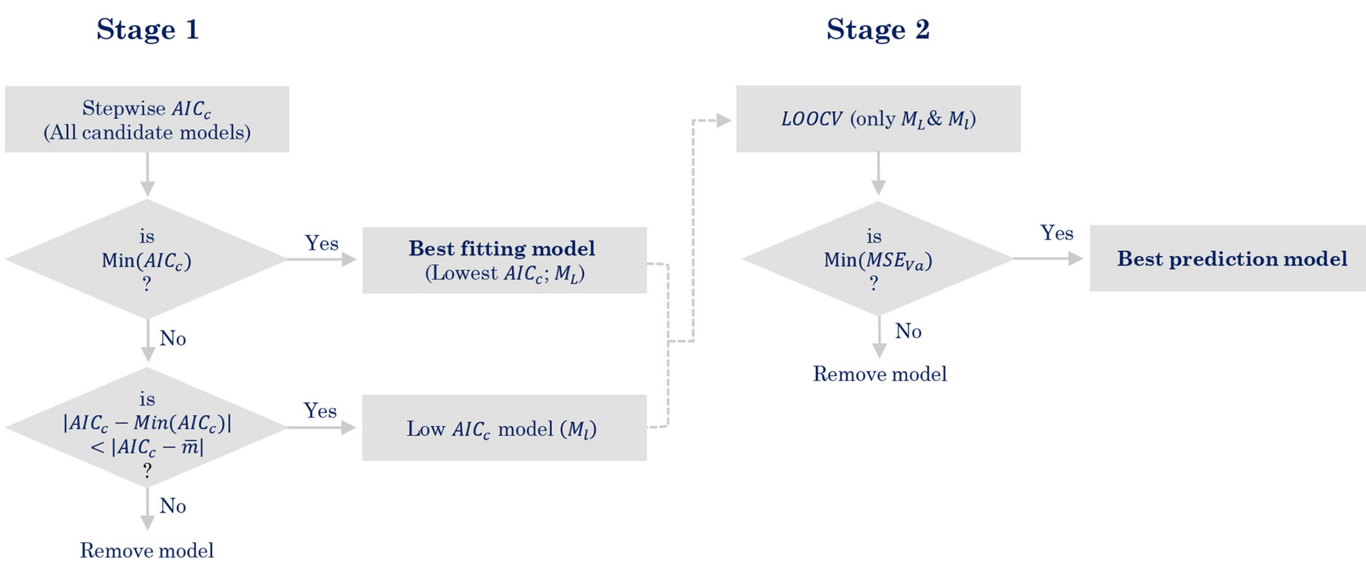

Fig 2. Two-stage selection of the best prediction model. $Min(AIC_c)$ and $Min(MSE_{Va})$ refer to the lowest value of AIC_c and lowest value of mean squared error of $LOOCV$ validation. $\bar{m}$ represents the average value of AIC_c for all candidate models.

https://doi.org/10.1371/journal.pgph.0000047.g002

lowest AIC_c model (M_L). In contrast, the models satisfying the condition

$$|AIC_c - Min(AIC_c)| < |AIC_c - \bar{m}| \tag{2}$$

were defined as low AIC_c models (M_l), where $\bar{m}$ is the average AIC_c of all candidate models. The reason we included both lowest and low AIC_c models is that we aim to identify the best prediction model using *LOOCV* among all the models with low AIC_c. In the second stage, the M_L and M_l models were compared using the mean squared errors obtained from leave-one-out cross-validation (*LOOCV*). The best model was the model with the minimum MSE_{Va} in *LOOCV*. In addition, the models were compared using $QAIC_c$ of which a lower value referred to a better predictivity of the model.

Model validation

Model validation was conducted using *LOOCV* to check how accurately the models can predict an independent dataset. If the difference between the predicted and observed value is minimum, we considered the predictive model is good. To perform *LOOCV*, the data for a specific test year were removed and the model was fitted based on the remaining data, which served as a training set. The fitted model was then used to predict the annual dengue cases for the test year. We repeated this procedure for all years from 2000 to 2018. Mean squared errors for the validation set MSE_{Va} and for the training set MSE_{Tr} were obtained by calculating the difference between predicted and observed numbers of annual dengue cases in the testing set and training set. Next, we checked the mean squared error ratio $\left(F = \frac{MSE_{Va}}{MSE_{Tr}}\right)$ [22], the ratio of the mean squared errors of the validation set and the training set.

We computed bootstrap confidence intervals for the predicted annual dengue cases in each year. To do this, we simulated 1,000 random samples from a Poisson distribution by considering *LOOCV*-estimated annual cases ($\hat{\lambda}$) as the parameter of the distribution. The random numbers were used to refit the model 1,000 times, giving the distribution of the estimated parameters. The lower and upper bounds of the 95% confidence intervals were calculated based on the 2.5% and 97.5% quantiles of the parameter distributions.

Assessment of the effects of climate factors

Interpretation of the estimated model coefficients for a generalized linear model is not as straightforward as it is for an ordinary linear regression model, as the dependent variable y is associated with a link function, such as a Poisson link [32]. Therefore, we calculated the marginal effect at the mean (MEM) to understand the effect of each of the predictor variables separately using the *R*-package *ggeffect* [33].

Results

Dengue cases in Bangladesh

The number of dengue cases exhibited a decreasing trend since the outbreak in 2000 until 2010 (Fig 3A). After 2010, the number began to increase until a drop in 2014, and then grow again till 2018. Over 5,000 infections were reported in 2000, 2002 and 2016. In contrast, in 2018, over 10,000 cases were reported. Dengue fever occurs primarily between July and November each year (Fig 3B). Therefore, monthly climate predictors were selected prior to July to predict the overall annual incidence.

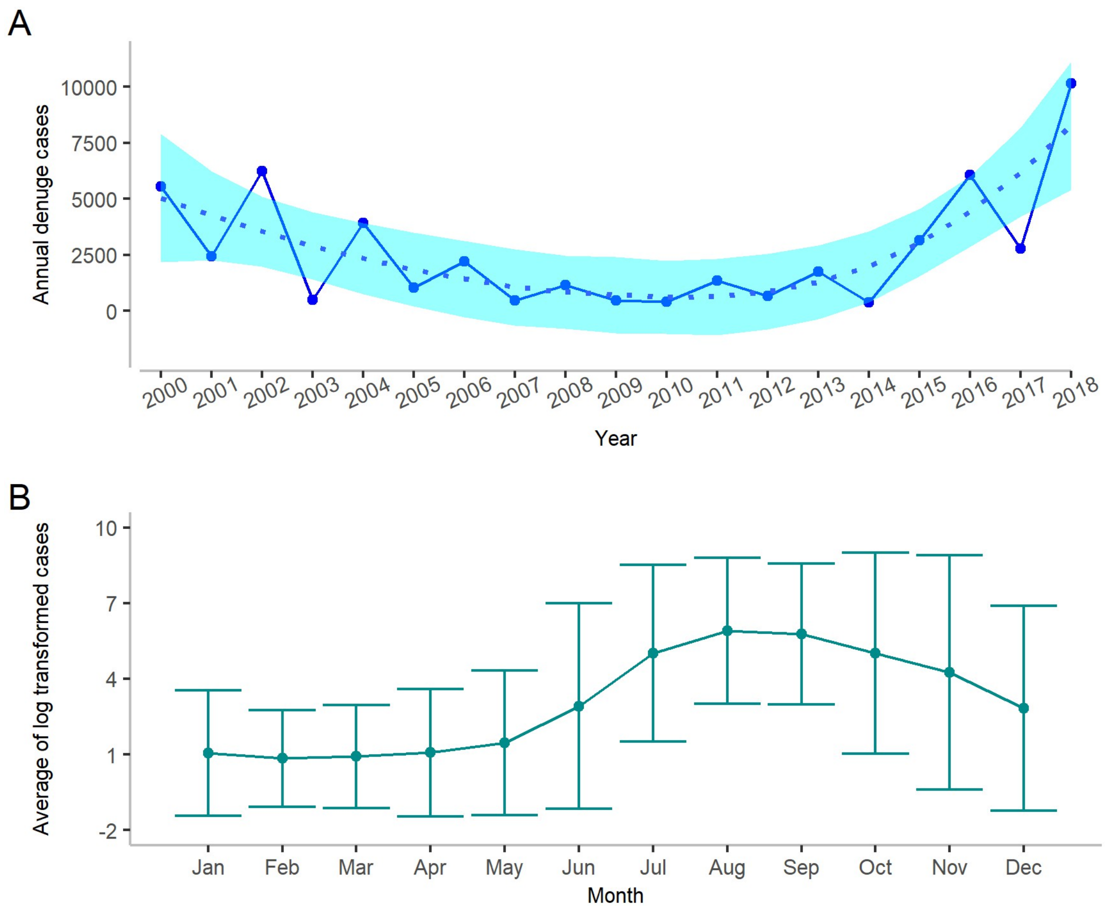

Fig 3. Dengue incidence in Bangladesh for different years and months. (**A**) Annual dengue case trends between 2000 and 2018. A LOESS smoothing function is used to obtain a smooth line to represent the trend over the years. The shaded region shows the pointwise 95% confidence interval. (**B**) Month-wise average dengue cases from January 2000 to December 2018. The shaded area represents the 95% confidence interval.

https://doi.org/10.1371/journal.pgph.0000047.g003

Climate variabilities

Different combinations of temperature (monthly average, maximum or minimum), rainfall (monthly total or maximum) and sunshine duration were used to predict annual dengue incidence. The monthly minimum temperature in Bangladesh increased after January and continued to increase until June/July, with the highest temperature of $26.4^{\circ}C$ measured in 2010 and 2014 (Fig 4). However, large variations in temperature were observed between March and June. Rainfall increased between April and October (S1 Fig). The highest amount of rainfall usually occurs between May and August, with low levels of rainfall recorded in January, February and December. Sunshine duration in January to April and in December exhibits a decreasing trend over the years, an indication of warmer winter (S1 Fig). Longer sunlight duration is mostly observed in the period from April to June.

Model selection and annual dengue prediction

To obtain an appropriate set of predictors for annual dengue prediction, we compared six models with different combinations of climate variables (Table 1). The best prediction model was determined using a two-stage model selection approach. In the first stage, Model 3 and Model 6, which belonged to either the low or the lowest AIC_c models, were chosen based on

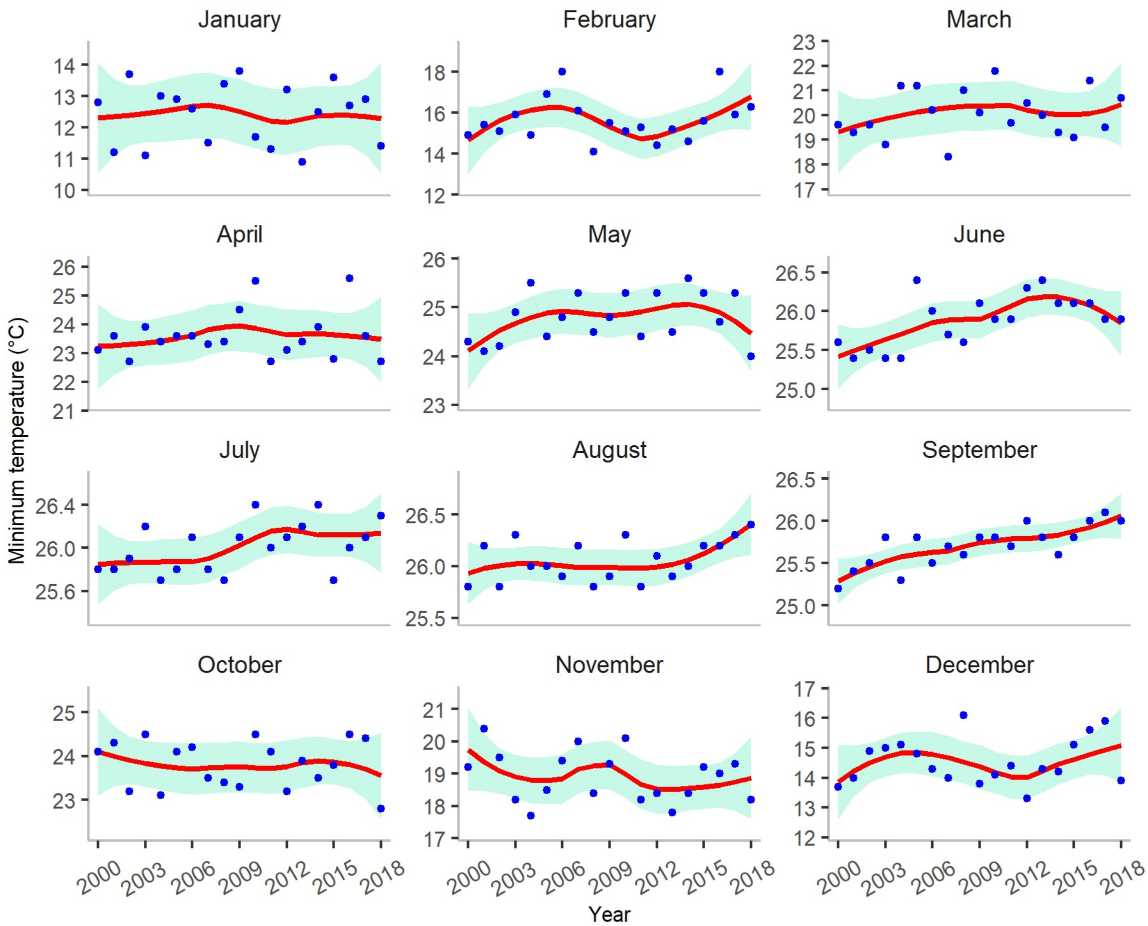

Fig 4. Minimum temperatures (in ˚*C*) in Bangladesh between January 2000 and December 2018. A LOESS smoothing function is used to obtain a smooth line to represent the trend over the years. The shaded region represents the 95% confidence interval. Dots represent the average minimum temperature of a given month for a particular year.

https://doi.org/10.1371/journal.pgph.0000047.g004

the criteria defined in Eq (2). For details of the stepwise AIC_c results, please refer to S2–S7 Tables. In the second stage, *LOOCV* was conducted for the selected models. After *LOOCV* was performed, Model 3 was identified as the best prediction model, with the lowest mean squared error for the validation set, compared with Model 6 (0.29 vs 0.31; see Table 2). Model 3 was

Table 1. Comparison of candidate models based on different evaluation metrics.

Model	Temperature	Rainfall	Sunshine	AIC_c	$QAIC_c$	R^2_V	$Adj - R^2_V$
Model 1	$ave.T_i$	$tot.R_i$	S_i	1,731	193	0.9903	0.9566
Model 2	$max.T_i$	$tot.R_i$	S_i	3,235	134	0.9899	0.9635
Model 3	$min.T_i$	$tot.R_i$	S_i	791	107	0.9951	0.9853
Model 4	$ave.T_i$	$max.R_i$	S_i	2,026	135	0.9932	0.9755
Model 5	$max.T_i$	$max.R_i$	S_i	3,913	191	0.9707	0.8683
Model 6	$min.T_i$	$max.R_i$	S_i	675	140	0.9961	0.9861

Models 1, 2 and 3 considered average, maximum and minimum monthly temperatures, respectively, along with monthly sunshine duration and monthly total rainfall. Models 4, 5 and 6 considered average, maximum and minimum monthly temperature, respectively, with monthly sunshine duration and maximum monthly rainfall. R^2_V and $Adj - R^2_V$ represent the pseudo-coefficient of determination and the adjusted coefficient of determination for the generalized linear model.

https://doi.org/10.1371/journal.pgph.0000047.t001

Table 2. The second stage of model selection.

Model	MSE_{Va}	MSE_{Tr}	F
Best fitting model (Model 6)	0.31	0.03	9.82
Best prediction model (Model 3)	0.29	0.05	5.73

Mean squared errors for the validation data set (MSE_{Va}) and the training data set (MSE_{Tr}) and their ratios $F = \frac{MSE_{Va}}{MSE_{Tr}}$ were used to select the best prediction model. These measures were calculated while performing *LOOCV* as given in S8 and S9 Tables.

https://doi.org/10.1371/journal.pgph.0000047.t002

also considered as the best fitting model while comparing the models using $QAIC_c$. Therefore, we used the best prediction model (Model 3) for further analysis of the impact of climate on dengue incidence. Note that the best fitting model (Model 6) also identified similar climate variables to the best prediction model. The best prediction model can be expressed as,

$$\begin{aligned}\log(\lambda_j) &= 57.6(\pm 11.039) + 0.41(\pm 0.097)T_{1j} + 0.36(\pm 0.095)T_{2j} \\ &+0.54(\pm 0.111)T_{3j} - 0.54(\pm 0.118)T_{4j} \\ &-0.72(\pm 0.228)T_{5j} - 1.34(\pm 0.241)T_{6j} \\ &-0.48(\pm 0.195)S_{4j} - 0.56(\pm 0.113)S_{5j} \\ &-0.01(\pm 0.012)R_{1j} + 0.02(\pm 0.007)R_{2j} \\ &+0.002(\pm 0.002)R_{4j} + 0.001(\pm 0.001)R_{6j}.\end{aligned} \tag{3}$$

In Eq (3), the estimated mean and corresponding standard deviation for each climate predictors were given. Using quasi-Poisson regression, the estimates for the minimum temperature and sunshine duration were statistically significant with p–value < 0.001 except sunshine duration in April. The coefficient of rainfall in February was also significant with p–value = 0.0393 (see S10 Table for details). We evaluated the performance of the model using an *LOOCV* technique. The best prediction model successfully predicted the largest outbreak in 2018 as well as smaller outbreaks in five different years (2003, 2006, 2010, 2012 and 2014) (Fig 5). The estimated number of annual dengue cases for other years fell within a narrow range of the 95% bootstrap confidence interval (S11 Table).

Marginal effect of climate predictors

To check the impact of each climate variable individually on annual dengue incidence, we further assessed the marginal effects of climate predictors. The optimal minimum temperature for mosquito population expansion is around 21–23$^\circ C$ (Fig 6). There was an upward trend of temperature from January until June. During this six-month period, the marginal effects of mean minimum temperature from January to March were positive. In contrast, the effects were negative from April to June. Thus, the results indicate that a turning point of marginal effects was located between 21 and 23$^\circ C$. Starting from 2,500 predicted cases with a temperature of 23$^\circ C$, the number of predicted cases gradually declined to below 1000 predicted cases in April if the temperature was increased by two degrees. The similar patterns were also evident in May and June.

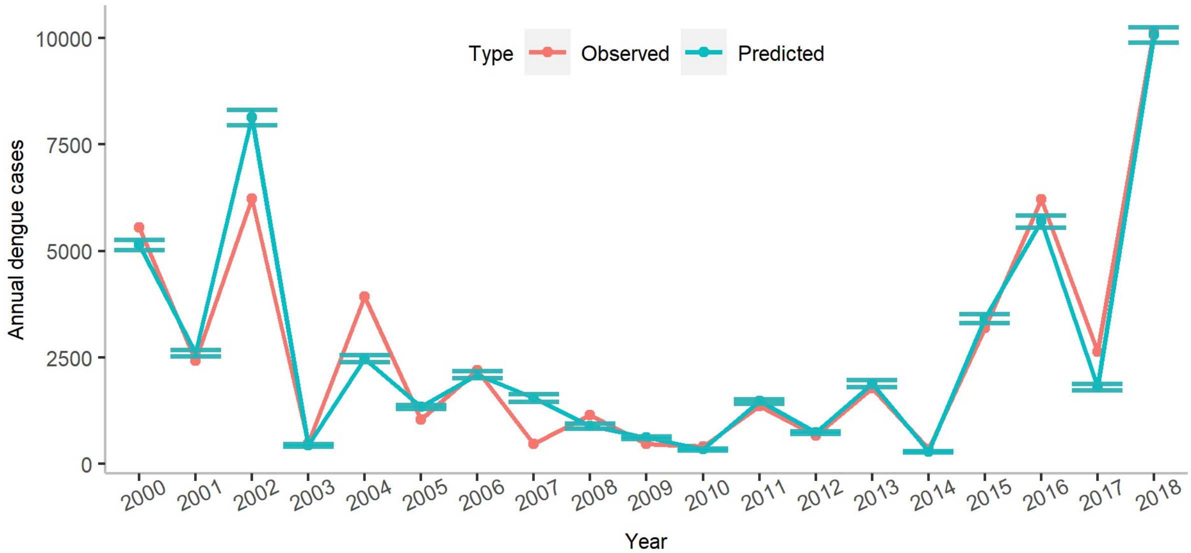

Fig 5. Comparison of observed and predicted annual dengue cases. The 95% confidence intervals were estimated using a bootstrap estimation technique and the *LOOCV* estimates.

https://doi.org/10.1371/journal.pgph.0000047.g005

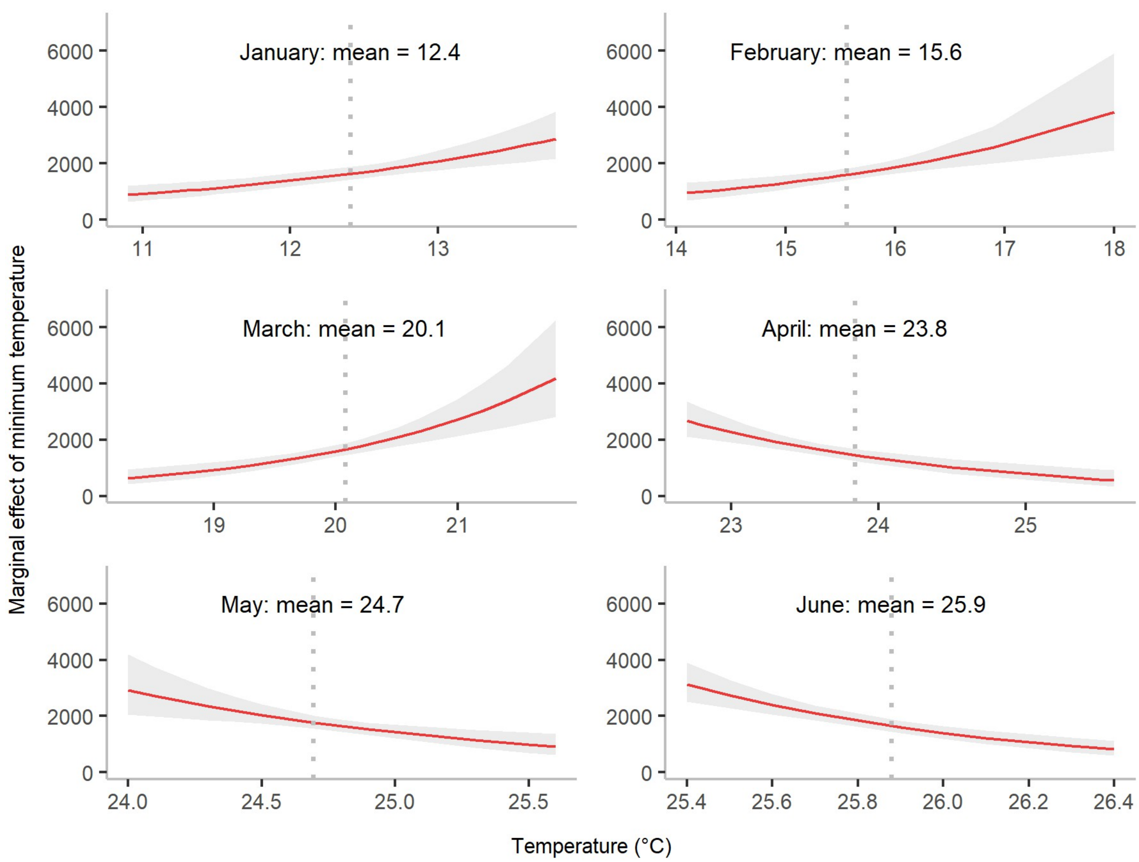

Fig 6. Marginal effect of minimum temperature from January to June on annual dengue cases. The shaded area denotes the 95% confidence interval of annual dengue cases at different values of minimum temperature, whereas the dotted vertical line represents mean minimum temperature for a month. The marginal effect here represents marginal effects at the mean (MEMs).

https://doi.org/10.1371/journal.pgph.0000047.g006

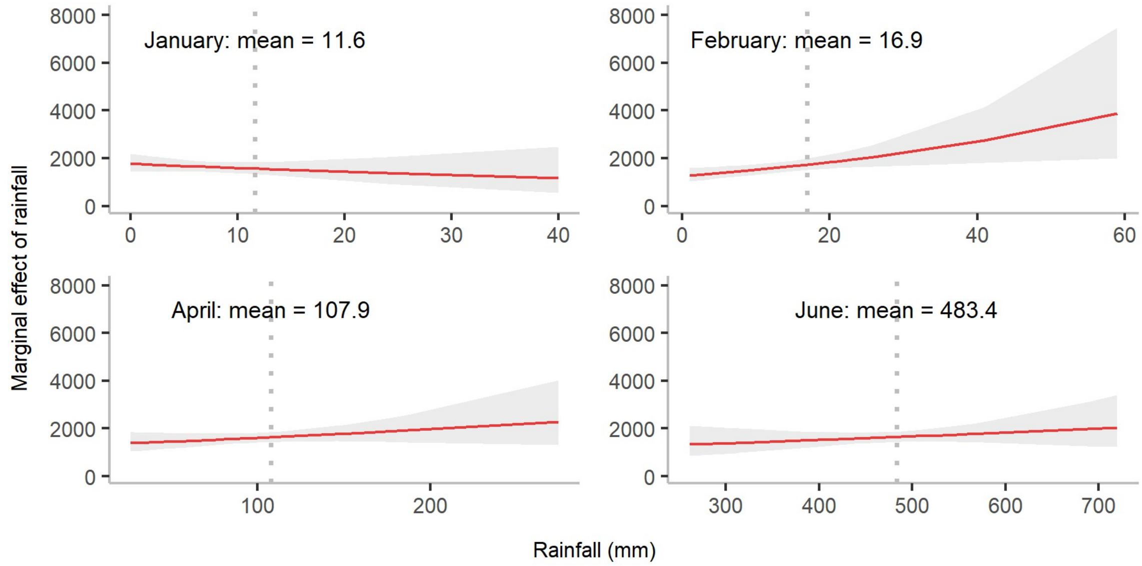

Fig 7. Marginal effect of total rainfall on annual dengue incidence in January, February, April and June. The shaded area denotes the 95% confidence interval of the expected number of annual dengue cases for different values of rainfall whereas the dotted vertical line represents mean rainfall for a month. The marginal effect here represents marginal effects at the mean (MEMs).

https://doi.org/10.1371/journal.pgph.0000047.g007

Rainfall also had different effects depending on the time. In February, April and June, rainfall had a positive relationship with dengue incidence (Fig 7). In contrast, rainfall in a cooler winter period (January) had a negative association with dengue incidence.

The mean sunshine duration in April and May were 7.4 and 6.5 hours, respectively and were negatively associated with annual dengue incidence (Fig 8). Comparing the magnitude of all climate predictors, minimum temperature in June (T6), sunshine duration in May (S5) and rainfall in February (R2) had the strongest effects on annual dengue incidence.

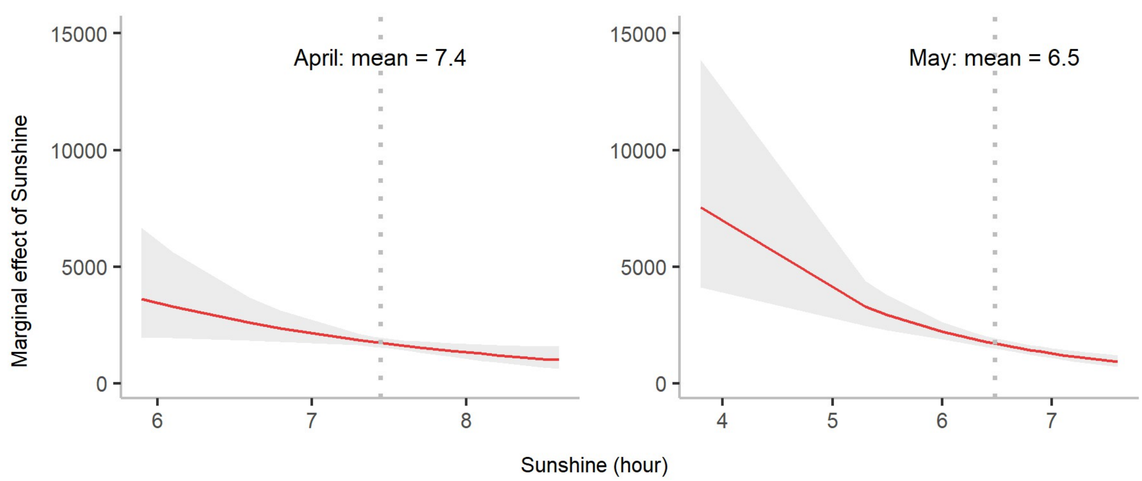

Fig 8. Marginal effect of sunshine duration on annual dengue incidence in April and May. The shaded area denotes the 95% confidence interval of the expected number of annual dengue cases for different values of sunshine duration, whereas the dotted vertical line represents mean sunshine duration for a month. The marginal effect here represents marginal effects at the mean (MEMs).

https://doi.org/10.1371/journal.pgph.0000047.g008

Discussion

Climate change poses a great threat to global health, particularly for subtropical and tropical climate regions due to the expansion of dengue fever. Since 2010, Bangladesh had an increasing total number of dengue cases, except in 2014, during each seasonal epidemic. In 2018, the recorded number of confirmed and suspected cases was more than 10,000, including 26 confirmed deaths [3, 34]. Although many studies have attempted to estimate dengue incidence in Bangladesh using climate data, most of these studies focused on data collected prior to 2010 [15–17, 35]. We developed a model to estimate the impact of various climate factors in the lead-up to dengue seasons, which typically occur during the monsoon season in Bangladesh. This is the first study to demonstrate that climate variability before dengue season can explain dengue expansions in Bangladesh in the past 20 years, suggesting that an early warning system can be built for this area.

Dengue fever has been increasing the health burden worldwide, including in South and Southeast Asian countries. Previous studies have been conducted in these regions to investigate the effects of climate change on dengue fever spreading [35–40]. In these studies, temperature, rainfall, and humidity were the commonly used climate predictors that can possibly influence dengue outbreaks. These studies showed an overall effect of particular climate variables throughout the years. In contrast, in our study, we explained that the impact of climate variables depends on the time of a year and is capable of predicting dengue cases before starting the peak season. It would help to mitigate any upcoming severe incidence in Bangladesh as it allows sufficient time to preparedness.

Our study suggests that some climate factors might exert opposing effects on the annual number of dengue cases, depending on the time of year. For example, minimum temperatures from January to March were positively associated with dengue cases, whereas a negative association was seen in the subsequent months from April to June, closer to the start of dengue season. This may be due to a complex dependency between the population dynamics of the dengue vector and the changing environment, such as the seasonal transition from winter to summer and the associated increasing temperatures. One study claimed that temperatures of 21.3–34$^{\circ}C$ are optimal for expansion of *Aedes aegypti* populations [21]. As the average daily minimum temperature is lower than 21$^{\circ}C$ before April and higher than 23$^{\circ}C$ during and after April, our results suggest an optimal daily minimum temperature in the range of 21–23$^{\circ}C$.

Total rainfall in a winter month (January) was found to have a negative relationship with dengue cases, whereas a positive relationship was seen for later months mainly in summer such as April and June. Rainfall is thought to have both beneficial and harmful effects on mosquito population growth. Rainfall can provide standing water for mosquito breeding. However, an excessive amount of rainfall (e.g., heavy rainfall during monsoons or cyclones) has been commonly thought to be able to disrupt potential mosquito habitats [41]. Our study has identified a negative relationship between rainfall and dengue incidence in January. This may be because during this period the number of adult mosquitoes is low, meaning that rainfall has a larger negative impact by flushing away mosquito eggs than a positive impact due to creating habitats required by adult mosquitoes. A similar pattern of a negative association of pre-dengue-season rainfall with dengue cases has been seen in recent studies [22–24].

Additionally, sunshine duration was found to be closely linked to mosquito-related activities, such as frequency of mosquito bites [42]. However, sunshine duration has not been included in any prediction models thus far. Therefore, an evaluation of the increasing dengue incidence since 2010, with respect to these climate variables, is warranted. A shorter duration of sunshine is more favorable for dengue transmission. In general, mosquitoes are more active in darker environments, and there is a greater chance of dengue being transmitted during periods of less sunshine due to the increasing frequency of mosquito bites [43]. The marginal effect of sunshine duration in Fig 8 reveals that the shorter the duration of sunlight, the higher the number of dengue cases, supporting the biological characteristics of mosquito activity as described by [44]. A 2-hour reduction in sunshine duration in April and May was predicted to result in a 3-fold increase in annual cases. These estimates are consistent with a previous study that found a negative association between sunshine duration and dengue incidence [43].

While this study has an implication, some limitations exist in this study. Meteorological data for 2019 were not disclosed at the time of this study; hence, 2019 data were excluded from the models. Secondly, the small number of outcome values (19 data points) and a relatively large number of predictor variables might lead to overfitting. Dengue prediction in 2002 and 2007 might be the consequence of it. Moreover, the epidemiological data includes both laboratory confirmed cases and probable case, hence, actual estimate might be affected due to under/over reporting biased. Because we aim to estimate the effects of climate variability before dengue season on annual incidence in this study we did not plan to include monthly predictors within dengue season.

In conclusion, our research offers a potential alert system by modeling annual dengue outbreaks before the season begins using climate variables. As an early warning system is required to improve public health and safety, the model we developed may improve disease control systems in Bangladesh. This research will aid our understanding of the effects of climate variability on dengue expansion, not only in Bangladesh but also in northern India and other Southeast Asian countries with similar climates and social-economic conditions.

Supporting information

S1 Fig. Sunshine duration (in hours) in Bangladesh between January 2000 and December 2018. A LOESS smoothing function is used to obtain a smooth line to represent the trend over the years. The shaded region represents the 95% confidence interval. Dots represent the average sunshine duration of a given month for a particular year.
(TIFF)

S2 Fig. Monthly total rainfall (in mm) in Bangladesh between January 2000 and December 2018. A LOESS smoothing function is used to obtain a smooth line to represent trend over years. The shaded region shows the 95% confidence interval. Dots represent the average rainfall of a given month for a particular year.
(TIFF)

S3 Fig. Scatter plot of sunshine duration in April and minimum temperature between January and June. The top left panel represents minimum temperature data in January, whereas the bottom right represents minimum temperature data in June. r denotes the correlation coefficients score and p is p–value from a correlation test. The line refers to the regression line, and the shaded region shows the 95% confidence interval. The points are intersecting values of minimum temperature and sunshine duration.
(TIFF)

S4 Fig. Scatter plot of sunshine duration in May and minimum temperature between January and June. The top left panel represents minimum temperature data in January, whereas the bottom right represents minimum temperature data in June. r denotes the correlation coefficients score and p is the p–value from a correlation test. The line refers to the regression line, and the shaded region shows the 95% confidence interval. The points are intersecting values of minimum temperature and sunshine duration.
(TIFF)

S1 Table. Variance inflation factor (VIF) for the predictors used in the best prediction model: Monthly minimum temperature, monthly sunshine duration, and monthly total rainfall.
(PDF)

S2 Table. (*Model 1*) Step-by-step forward selection results of the generalized Poisson regression model for each step based on AIC_c. $ave.T_i$, S_i and $tot.R_i$ represent mean temperature, sunshine duration and total rainfall in the i^{th} month. For each of the variables included in the model, the corresponding AIC_c, the leave-one-out mean squared error for the validation set (MSE_{Va}), the leave-one-out mean squared error for the training set (MSE_{Tr}), and the mean squared error ratio $\left(F = \frac{MSE_{va}}{MSE_{Tr}}\right)$ were calculated.
(PDF)

S3 Table. (*Model 2*) Step-by-step forward selection results of the generalized Poisson regression model for each step based on AIC_c. $max.T_i$, S_i and $tot.R_i$ represent maximum temperature, sunshine duration and total rainfall in the i^{th} month. For each of the variables included in the model, the corresponding AIC_c, the leave-one-out mean squared error for the validation set (MSE_{Va}), the leave-one-out mean squared error for the training set (MSE_{Tr}), and the mean squared error ratio $\left(F = \frac{MSE_{va}}{MSE_{Tr}}\right)$ were calculated.
(PDF)

S4 Table. (*Model 3*) Step-by-step forward selection results of the generalized Poisson regression model for each step based on AIC_c. T_i, S_i and R_i represent minimum temperature, sunshine duration and total rainfall in the i^{th} month. For each of the variables included in the model, the corresponding AIC_c, the leave-one-out mean squared error for the validation set (MSE_{Va}), the leave-one-out mean squared error for the training set (MSE_{Tr}), and the mean squared error ratio $\left(F = \frac{MSE_{va}}{MSE_{Tr}}\right)$ were calculated.
(PDF)

S5 Table. (*Model 4*) Step-by-step forward selection results of the generalized Poisson regression model for each step based on AIC_c. $ave.T_i$, S_i and $max.R_i$ represent mean temperature, sunshine duration and maximum rainfall in the i^{th} month. For each of the variables included in the model, the corresponding AIC_c, the leave-one-out mean squared error for the validation set (MSE_{Va}), the leave-one-out mean squared error for the training set (MSE_{Tr}), and the mean squared error ratio $\left(F = \frac{MSE_{va}}{MSE_{Tr}}\right)$ were calculated.
(PDF)

S6 Table. (*Model 5*) Step-by-step forward selection results of the generalized Poisson regression model for each step based on AIC_c. $max.T_i$, S_i and $max.R_i$ represent maximum temperature, sunshine duration and maximum rainfall in the i^{th} month. For each of the variables included in the model, the corresponding AIC_c, the leave-one-out mean squared error for the validation set (MSE_{Va}), the leave-one-out mean squared error for the training set (MSE_{Tr}), and the mean squared error ratio $\left(F = \frac{MSE_{va}}{MSE_{Tr}}\right)$ were calculated.
(PDF)

S7 Table. (*Model 6*) Step-by-step forward selection results of the generalized Poisson regression model for each step based on AIC_c. $min.T_i$, S_i and $max.R_i$ represent minimum temperature, sunshine duration and maximum rainfall in the i^{th} month. For each of the

variables included in the model, the corresponding AIC_c, the leave-one-out mean squared error for the validation set (MSE_{Va}), the leave-one-out mean squared error for the training set (MSE_{Tr}), and the mean squared error ratio $\left(F = \frac{MSE_{va}}{MSE_{Tr}}\right)$ were calculated.
(PDF)

S8 Table. Leave-one-out cross-validation (LOOCV) results for Model 6. Achieved by omitting the j^{th} year in the j^{th} iteration, where $j = 1, \ldots, 19$, and $j = 1$ indicate the year 2000, $j = 2$ indicates 2001, etc. Italic font denotes the predicted annual dengue cases when the j^{th} year is removed.
(PDF)

S9 Table. Leave-one-out cross-validation (LOOCV) results for Model 3. Achieved by omitting the j^{th} year in the j^{th} iteration, where $j = 1, \ldots, 19$, and $j = 1$ indicates the year 2000, $j = 2$ indicates 2001, etc. Italic font denotes the predicted annual dengue cases when the j^{th} year is removed.
(PDF)

S10 Table. Parameter estimates of the best prediction model based on quasi Poisson regression. SD represents the standard deviations of the estimate of each predictor. Asterisks in the p–value indicates that the predictors are significant with certain levels (i.e. $^{**} = 0.001$; $^{*} = 0.01$).
(PDF)

S11 Table. Comparison between observed and predicted annual dengue cases in Bangladesh between 2000 and 2018. The lower and upper boundaries represent the lower and upper limit of the 95% bootstrap confidence interval, respectively, for the predicted value.
(PDF)

S12 Table. Comparison of negative binomial regression models based on AIC_c.
(PDF)

S13 Table. Comparison of the validation results among the best fitting models in negative binomial regression.
(PDF)

S1 Data.
(CSV)

S1 Text.
(DOCX)

Acknowledgments

We thank Dr. Iqbal Ansary Khan, Principal Scientific Officer, and Head of Medical Social Science, Institute of Epidemiology, Disease Control and Research, Directorate General of Health Services, Dhaka, for generously providing dengue surveillance data. The authors are indebted to City University of Hong Kong for providing excellent research facilities.

Author Contributions

Conceptualization: M. Pear Hossain, Hsiang-Yu Yuan.

Data curation: M. Pear Hossain.

Formal analysis: M. Pear Hossain.

Funding acquisition: Hsiang-Yu Yuan.

Methodology: M. Pear Hossain, Wen Zhou.

Software: M. Pear Hossain.

Supervision: Hsiang-Yu Yuan.

Visualization: M. Pear Hossain.

Writing – original draft: M. Pear Hossain, Hsiang-Yu Yuan.

Writing – review & editing: M. Pear Hossain, Wen Zhou, Chao Ren, John Marshall, Hsiang-Yu Yuan.

References

1. Guo C, Zhou Z, Wen Z, Liu Y, Zeng C, Xiao D, et al. Global Epidemiology of Dengue Outbreaks in 1990-2015: A Systematic Review and Meta-Analysis. Frontiers in Cellular and Infection Microbiology. 2017; 7:317. https://doi.org/10.3389/fcimb.2017.00317 PMID: 28748176
2. Ferreira GLC. Global Dengue Epidemiology Trend. Revista do Instituto de Medicina Tropical de São Paulo. 2012; 54:5–6. https://doi.org/10.1590/S0036-46652012000700003
3. World Health Organization. Dengue and severe dengue; 2020. https://www.who.int/news-room/fact-sheets/detail/dengue-and-severe-dengue.
4. Stanaway JD, Shepard DS, Undurraga EA, Halasa YA, Coffeng LE, Brady OJ, et al. The global burden of dengue: an analysis from the Global Burden of Disease Study 2013. The Lancet Infectious Diseases. 2016; 16(6):712–723. https://doi.org/10.1016/S1473-3099(16)00026-8 PMID: 26874619
5. Pasin CI, Elizabeth Halloran M, Gilbert PB, Langevin E, Leon Ochiai R, Pitisuttithum P, et al. Periods of high dengue transmission defined by rainfall do not impact efficacy of dengue vaccine in regions of endemic disease. Plos One. 2018. https://doi.org/10.1371/journal.pone.0207878
6. Loo YY, Billa L, Singh A. Effect of climate change on seasonal monsoon in Asia and its impact on the variability of monsoon rainfall in Southeast Asia. Geoscience Frontiers. 2015; 6(6):817–823. https://doi.org/10.1016/j.gsf.2014.02.009
7. Louis VR, Montenegro Quiñonez CA, Kusumawathie P, Palihawadana P, Janaki S, Tozan Y, et al. Characteristics of and factors associated with dengue vector breeding sites in the City of Colombo, Sri Lanka. Pathogens and Global Health. 2016; 110(2):79–86. https://doi.org/10.1080/20477724.2016.1175158 PMID: 27241954
8. Getachew D, Tekie H, Gebre-Michael T, Balkew M, Mesfin A. Breeding sites of Aedes aegypti: Potential dengue vectors in Dawa, east Ethiopia. Interdisciplinary Perspectives on Infectious Diseases. 2015; 2015. https://doi.org/10.1155/2015/706276 PMID: 26435712
9. Evelyn N, Murray A, Quam MB, Wilder-Smith A. Epidemiology of dengue: past, present and future prospects. Clinical Epidemiology. 2013. https://doi.org/10.2147/CLEP.S34440
10. World Health Organization. Ten threats to global health in 2019; 2020. https://www.who.int/vietnam/news/feature-stories/detail/ten-threats-to-global-health-in-2019.
11. Russell PK, Buescher EL, McCown JM, Ordõnez J. Recovery of dengue viruses from patients during epidemics in Puerto Rico and East Pakistan. The American Journal of Tropical Medicine and Hygiene. 1966; 15(4):573–579. https://doi.org/10.4269/ajtmh.1966.15.573 PMID: 4957424
12. Rahman M, Rahman K, Siddque AK, Shoma S, Kamal AHM, Ali KS, et al. First Outbreak of Dengue Hemorrhagic Fever, Bangladesh. Emerging Infectious Diseases. 2002; 8(7). https://doi.org/10.3201/eid0807.010398 PMID: 12095447
13. Institute of Epidemiology, Disease Control and Research. Web-based Dengue Surveillance; 2020. https://www.iedcr.gov.bd/index.php/surveillance.
14. Hsan K, Hossain MM, Sarwar MS, Wilder-Smith A, Gozal D. Unprecedented rise in dengue outbreaks in Bangladesh; 2019.
15. Sharmin S, Glass K, Viennet E, Harley D. Geostatistical mapping of the seasonal spread of under-reported dengue cases in Bangladesh. PLoS Neglected Tropical Diseases. 2018; 12(11):1–13. https://doi.org/10.1371/journal.pntd.0006947 PMID: 30439942

16. Banu S, Hu W, Guo Y, Hurst C, Tong S. Projecting the impact of climate change on dengue transmission in Dhaka, Bangladesh. Environment International. 2014; 63:137–142. https://doi.org/10.1016/j.envint.2013.11.002 PMID: 24291765

17. Hashizume M, Dewan AM, Sunahara T, Rahman MZ, Yamamoto T. Hydroclimatological variability and dengue transmission in Dhaka, Bangladesh: A time-series study. BMC Infectious Diseases. 2012; 12. https://doi.org/10.1186/1471-2334-12-98 PMID: 22530873

18. Lai YH. The climatic factors affecting dengue fever outbreaks in southern Taiwan: An application of symbolic data analysis. BioMedical Engineering Online. 2018; 17(s2):1–14. https://doi.org/10.1186/s12938-018-0575-4 PMID: 30396346

19. Lu L, Lin H, Tian L, Yang W, Sun J, Liu Q. Time series analysis of dengue fever and weather in Guangzhou, China. BMC Public Health. 2009; 9:1–5. https://doi.org/10.1186/1471-2458-9-395 PMID: 19860867

20. Gu H, Leung RKK, Jing Q, Zhang W, Yang Z, Lu J, et al. Meteorological factors for dengue fever control and prevention in South China. International Journal of Environmental Research and Public Health. 2016; 13(9):1–12. https://doi.org/10.3390/ijerph13090867 PMID: 27589777

21. Ryan Id SJ, Carlson CJ, Mordecai EA, Johnson LR. Global expansion and redistribution of Aedes-borne virus transmission risk with climate change. PLOS Neglected Tropical Diseases. 2019; https://doi.org/10.1371/journal.pntd.0007213.

22. Yuan HY, Wen TH, Kung YH, Tsou HH, Chen CH, Chen LW, et al. Prediction of annual dengue incidence by hydro-climatic extremes for southern Taiwan. International Journal of Biometeorology. 2019; 63(2):259–268. https://doi.org/10.1007/s00484-018-01659-w PMID: 30680621

23. Yuan HY, Liang J, Lin PS, Sucipto K, Tsegaye MM, Wen TH, et al. The effects of seasonal climate variability on dengue annual incidence in Hong Kong: A modelling study. Scientific Reports. 2020; 10(1):1–10. https://doi.org/10.1038/s41598-020-60309-7

24. Lowe R, Gasparrini A, Van Meerbeeck CJ, Lippi CA, Mahon R, Trotman AR, et al. Nonlinear and delayed impacts of climate on dengue risk in Barbados: A modelling study. PLOS Medicine. 2018; 15 (7). https://doi.org/10.1371/journal.pmed.1002613 PMID: 30016319

25. World Health Organization. South-East Asia; 2020. https://www.who.int/southeastasia.

26. Worldatlas. Where Is Bangladesh?; 2020. https://www.worldatlas.com/as/bd/where-is-bangladesh.html.

27. Banglapedia. Rainfall; 2020. http://en.banglapedia.org/index.php?title=Rainfall.

28. Banglapedia. Bangladesh Geography; 2020. http://en.banglapedia.org/index.php?title=Bangladesh_Geography.

29. Directorate General of Health Services. National Guideline for Clinical Management of Dengue Syndrome. 4th ed. Dhaka: National Malaria Elimination & Aedes Transmitted Diseases Control Program Disease Control Unit; 2020.

30. Kim HJ, Cavanaugh JE, Dallas TA, Foré SA. Model selection criteria for overdispersed data and their application to the characterization of a host-parasite relationship. Environmental and Ecological Statistics. 2014; 21(2):329–350. https://doi.org/10.1007/s10651-013-0257-0

31. Hurvich CM, Tsai CL. Regression and Time Series Model Selection in Small Samples. Biometrika. 1989; 76(2):297. https://doi.org/10.1093/biomet/76.2.297

32. Fernihough A. mfx: Marginal Effects, Odds Ratios and Incidence Rate Ratios for GLMs; 2019. https://CRAN.R-project.org/package=mfx.

33. Lüdecke D. ggeffects: Tidy Data Frames of Marginal Effects from Regression Models. Journal of Open Source Software. 2018; 3(26):772. https://doi.org/10.21105/joss.00772

34. Directorate General of Health Services. Daily Dengue Status Report; 2020. https://dghs.gov.bd/index.php/bd/home/5200-daily-dengue-status-report.

35. Karim MN, Munshi SU, Anwar N, Alam MS. Climatic factors influencing dengue cases in Dhaka city: a model for dengue prediction. The Indian journal of medical research. 2012; 136(1):32–39. PMID: 22885261

36. Dom NC, Hassan AA, Latif ZA, Ismail R. Generating temporal model using climate variables for the prediction of dengue cases in Subang Jaya, Malaysia. Asian Pacific journal of tropical disease. 2013; 3 (5):352–361. https://doi.org/10.1016/S2222-1808(13)60084-5

37. Pinto E, Coelho M, Oliver L, Massad E. The influence of climate variables on dengue in Singapore. International journal of environmental health research. 2011; 21(6):415–426. https://doi.org/10.1080/09603123.2011.572279 PMID: 21557124

38. Polwiang S. The time series seasonal patterns of dengue fever and associated weather variables in Bangkok (2003-2017). BMC Infectious Diseases. 2020; 20(1):1–10. https://doi.org/10.1186/s12879-020-4902-6

39. Aswi A, Cramb S, Duncan E, Hu W, White G, Mengersen K. Climate variability and dengue fever in Makassar, Indonesia: Bayesian spatio-temporal modelling. Spatial and Spatio-temporal Epidemiology. 2020; 33:1–8. https://doi.org/10.1016/j.sste.2020.100335 PMID: 32370940

40. 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; 20(1):1–10. https://doi.org/10.1186/s12889-020-08846-8

41. Benedum CM, Seidahmed OME, Eltahir EAB, Markuzon N. Statistical modeling of the effect of rainfall flushing on dengue transmission in Singapore. PLOS Neglected Tropical Diseases. 2018; 12(12). https://doi.org/10.1371/journal.pntd.0006935 PMID: 30521523

42. Vu HH, Okumura J, Hashizume M, Tran DN, Yamamoto T. Regional differences in the growing incidence of dengue fever in Vietnam explained by weather variability. Tropical Medicine and Health. 2014; 42(1):25–33. https://doi.org/10.2149/tmh.2013-24 PMID: 24808744

43. Pham HV, Doan HTM, Phan TTT, Tran Minh NN. Ecological factors associated with dengue fever in a Central Highlands province, Vietnam. BMC Infectious Diseases. 2011; 11(1):172. https://doi.org/10.1186/1471-2334-11-172 PMID: 21679398

44. Kim YM, Park JW, Cheong HK. Estimated effect of climatic variables on the transmission of Plasmodium vivax malaria in the Republic of Korea. Environmental Health Perspectives. 2012; 120(9):1314–1319. https://doi.org/10.1289/ehp.1104577 PMID: 22711788