**2.Materials and Methods**

2.1. Dengue Cases and Meteorological Factors

Confirmed positive and suspected cases of dengue were extracted from the Bangladesh Rohingya Emergency Response Bulletin, a joint effort of the WHO and the Ministry of Health and Family Welfare Bangladesh. The data encompassed patients from Ukhiya and Teknaf regions of Cox’s bazar ,Bangladesh. For this study, we utilized the weekly dengue new cases reported between January 1, 2021, and August 5, 2023, available on the WHO website (<https://www.who.int/bangladesh/emergencies/Rohingyacrisis/ewars>).

To enrich our analysis, we incorporated meteorological information sourced from NASA's website (NASA, 2023) (<https://www.nasa.gov/>). The study specifically encompassed averaged weekly Earth Skin Temperature (℃), Specific Humidity at 2 Meters (g/kg), Relative Humidity at 2 Meters (%), Temperature at 2 Meters, Precipitation Correlated (mm/day), and Wind Speed at 2 Meters (m/s) as variables for our time series analysis.

2.2. Statistical Time Series Models

The Simple Exponential Smoothing Model (SES), Auto-Regressive Integrated Moving Average Model (ARIMA), Seasonal Auto-Regressive Integrated Moving Average Model (SARIMA), and GA time series models were used in this work to forecast recent Dengue cases. Additionally, Auto-Regressive Integrated Moving Average with Explanatory Variables (ARIMAX), Generalized Additive Model (GA), and Generalized Linear Mixed Models (GLM) time series models were used in this work to evaluate the correlations between meteorological parameters and Dengue cases. The SES model is utilized as a baseline to assess the prediction accuracy of other models in this investigation. All models were used to forecast new DENV cases. (ref.: ijerph-20-05152)

2.2.1. Simple Exponential Smoothing Model (SES)

Simple exponential smoothing is a short-range forecasting method that assumes a reasonably stable mean in the data with no trend (consistent growth or decline). SES is important because it is a relatively simple yet powerful and versatile technique for forecasting time series data. It is one of the most popular forecasting methods that uses weighted moving average of past data as the basis for a forecast. The SES model is given by the model equation:

Where β(t) takes a constant at the time t and may change slowly over the time; is a random variable and is used to describe the effect of stochastic fluctuation.

Let an observed time series be y1, y2, . . ., yn. In any case, in this simple model, to predict yt is merely to predict (estimate) β. To estimate, it makes sense to use all the past observations, but due to declining correlation as you go back into the past, to down-weight older observations. Formally, the simple exponential smoothing equation takes the form of

Ft+1 = α yt + (1−α)Ft ,

Where yt is the actual, known series value at the time t; Ft is the forecast value of the variable Y at the time t; Ft+1 is the forecast value at the time t + 1; α is the smoothing constant. (10.2478@v10198-012-0034-2 (1)).

2. 2.2. Auto-Regressive Integrated Moving Average Model (ARIMA)

ARIMA models are generally used to analyze time series data for better understanding and forecasting. If a Durbin-Watson statistic reveals that there is autocorrelation of sequential data points, analysis of variance and regression results will be invalid and possibly misleading. Such data sets may be analyzed by time series methodologies such as autoregressive integrated moving average (ARIMA) modeling. (ARIMA-2). The ARIMA model is a combination of three components: autoregression (AR), differencing (I), and moving average (MA). The AR component models the relationship between an observation and several lagged observations. The I component is used to make the time series stationary, which means that the statistical properties of the time series do not change over time. The MA component models the relationship between an observation and a residual error from a moving average model applied to lagged observations (ARIMA-4).

The AR(p) model refers to the autoregressive model of order p.

The AR(p) model is written:

yt = c+a1yt-1 + ………..+apyt-p +ut

where a1 ,a2,….,ap are parameters;

c is a constant and the random variable ut is white noise.

The MA(q) model refers to the moving average model of order q, The MA(q) model is written:

yt =u+ut +m1 ut-1 + …….+ mqut-q

Where the m1,m2,……,mq are the parameters of the model; u is the expectation of yt ( often assumed to equal 0); ut, ut-1, ut-2 , .., ut-q are white noise error terms.

The ARMA(p, q) model refers to the model with p autoregressive terms and q moving-average terms. This model contains the AR(p) and MA(q) models, and is written:

yt =c+a1yt-1 +…..+ apyt-p +ut+ ……. + mqut-q

When AR(P), MA(q) and ARMA(p,q) are applied in some cases where data show evidence of non-stationarity, an initial differencing step should be applied to reduce the non-stationarity, namely an ARIMA model. Non-seasonal ARIMA models are generally denoted ARIMA(p,q,d), where p is the order of the AR model, d is the degree of differencing, and q is the order of the MA model.(ARIMA-4).

2.2.3. Auto-Regressive Integrated Moving Average with Explanatory Variables (ARIMAX)

Recently, ARIMA has been studied by many researchers who used time series. However, when using ARIMA model, only one variable can be used, so it is not adequate to express real problems. Complex problems always need more the one variable in order to explain problems effectively. Therefore, it is necessary to build a multivariate ARIMAX model (Fan et al. 2009; Jalalkamali et al. 2015). The ARIMAX model (Bierens, 1987) is a generalization of the ARIMA model, which is capable of incorporating an external input variable (X)). ARIMAX model consists of four parts. These parts are Auto Regressive (AR), Integrated (I), Moving Average (MA), and Exogenous Variable (X) (Sutthichaimethee and Ariyasajjakorn, 2017 (ARIMAX-3). Nonlinear least square method employed to estimate the parameters of ARIMAX model. (Ref. ARIMAX).

2.2.4. Seasonal Auto-Regressive Integrated Moving Average Model (SARIMA)

The seasonal auto-regressive integrated moving average (SARIMA) model is a time series forecasting method proposed by Box and Jenkins in the 1970s. The general form of the SARIMA model is as follows:

(p,d,q)(P,D,Q)S

where p and q are the orders of auto-regressive (AR) and moving average (MA), respectively, d is the order of the differences, P, D and Q are the corresponding seasonal orders, and S represents the steps of the seasonal differences. (Ref.SARIMA-2).

The SARIMA considers the seasonal effect in the data series. If there is any seasonal or cyclic pattern in the series, the SARIMA model would be more appropriate to be used. There are other models which have similar concept with the ARMA model. One example would be the autoregressive conditional heteroscedastic (ARCH) model. However, the more commonly used models from the ARMA family are the ARMA model itself, the ARIMA model and the SARIMA model. (SARIMA-5).

2.2.5. Generalized Additive Model (GA)

The widely used generalized additive models (GAM) method is a flexible and effective technique for conducting nonlinear regression analysis in time-series studies of the health effects of air pollution. When the data to which the GAM are being applied have two characteristics—1) the estimated regression coefficients are small and 2) there exist confounding factors that are modeled using at least two nonparametric smooth functions. (Ref. GAM). The GAM, or generalized additive model, is applied for analyzing the interaction of climatic factors with Dengue cases already being used for various research purposes. This time series model is more eminent and user-friendly, which is preferable for analyzing national morbidity, mortality, and air pollution studies. (Ref. ijerph)

2.2.6. Generalized Linear Mixed Models (GLM)

The generalized linear mixed model (GLMM) generalizes the standard linear model in three ways: accommodation of non-normally distributed responses, specification of a possibly non-linear link between the mean of the response and the predictors, and allowance for some forms of correlation in the data. As such, GLMMs have broad utility and are of great practical importance. Two special cases of the GLMM are the linear mixed model (LMM) and the generalized linear model (GLM). Generalized linear models (GLM; McCullagh and Nelder 1989) provide an extension of linear models which relaxes the assumptions of normality, constant error variance and a linear relationship between the covariate effects and the mean.

2.3. Statistical Analysis

This study employed five different statistical time series models to analyze trends and forecast future Dengue cases. Additionally, it used three time series models to assess the correlation between Dengue cases and meteorological conditions. The investigation also made use of Spearman's rank correlation coefficients. To identify seasonal variations, a Poisson regression model was applied during three major seasons in Bangladesh. The choice of the Poisson distribution was based on the characteristics of the data, which involve counting observations as whole numbers, independence of event occurrences, and consistent time intervals for each participant.