Rwanda Journal of Engineering, Science, Technology and Environment, Volume 7, Issue1, March 2025
eISSN: 2617-233X |print ISSN: 2617-2321

# The Effect of Meteorological Factors on Extreme COVID-19 Infection in Rwanda: The Generalized Additive Extreme Value Modeling approach

**Innocent Ngaruye*, Joseph Nzabanita[1], Emmanuel Masabo[1,2], Marcel Gahamanyi[1] Bosco Nyandwi[1], Justine Dushimirimana[1], Denis Ndanguza[1], Lydie Mpinganzima[1], Célestin Kurujyibwami[1], Leon Fidele Uwimbabazi Ruganzu[1], Wellars Banzi[1], Venuste Nyagahakwa[1], Solange Mukeshimana[1], Jean Pierre Muhirwa[1] Jean Pierre Ngendahayo[1], Jean Paul Nsabimana[1], Jean de Dieu Niyigena[1], Eric Rutaganda[3], Jeanne Uwonkunda[4], Japhet Niyobuhungiro[5], Isambi Sailon Mbalawata[6], Betty Kivumbi Nannyonga[7] and Jean Marie Ntaganda[1]**

[1] College of Science and Technology, University of Rwanda, Rwanda
[2] African Centre of Excellence in Data Science, University of Rwanda, Rwanda
[3] College of Medicine and Health Sciences, University of Rwanda, Rwanda
[4] Great Mind Career for Development Group, Rwanda
[5] National Council for Science and Technology, Rwanda
[6] African Institute of Mathematical Sciences, Rwanda
[7] College of Natural Sciences, Makerere University, Uganda

Corresponding author: i.ngaruye@ur.ac.rw or ingaruye@gmail.com

**Abstract**

The novel human coronavirus disease, COVID-19, was first identified in China in 2019 and has since spread throughout the world becoming a global pandemic of great concern. High daily new cases have brought a heavy burden on health facilities and health workers helping patients and fighting the spread of this pandemic. Understanding the behavior of extreme cases of COVID-19 and associated factors is crucial to devise strategies to flatten the pandemic curve. This study used generalized additive modeling and extreme value theory approaches to analyze weekly maximum positive cases of COVID-19 together with three climate covariates (temperature, rainfall, and solar radiation) with the purpose to evaluate the predictive power of climate factors on extreme COVID-19 cases. According to the findings, a Generalized Extreme Value distribution with a constant location parameter, a

 http://dx.doi.org/10.4314/rjeste.v7i1.8

linear model for the shape parameter with rainfall as a predictor, and a non-linear model for the scale parameter with temperature and rainfall as predictors fits the weekly maximum positive cases the best. As a result, both temperature and rainfall have a significant effect on the spread of the COVID-19 pandemic. The findings of this study make a significant contribution to the existing knowledge about the COVID-19 pandemic.

**Keywords:** COVID-19, Climate factors, Extreme positive cases, Vector Generalized Linear models

# 1 Introduction

Since the appearance of its first new case in Wuhan (Chine), the entire globe is preoccupied with finding and making decisions about measures to fight against the COVID-19 pandemic. The control of the pandemic at different times has an important impact on the spread of the disease. Communities all over the world have been doing their best to protect their respective population applying various measures such as the social distancing in public and cleaning hands regularly, wearing face masks, stopping local and international transport, locking down cities, quarantine of contacts and banning gatherings [1–3]. In addition, depending on how the pandemic is spreading, Governments increase the Intensive Care Units for treatment of patients [4].

To aid medical practitioners, policymakers, and governing bodies in conducting intervention steps to manage the pandemic, modeling techniques have been widely used to anticipate COVID-19 spread. Due to the aforementioned continuing alarming situation, it is necessary to study the dynamics of COVID-19 to identify various factors related particularly to extreme positive cases. Particularly, in order to take appropriate measures of controlling the pandemic in situations where we observe extreme new cases, there is a need to analyze the main causes of these situations and associated determinants including, but not limited to climate factors, treatment the facility, chronic diseases, and existing Government preventive measures.

Several studies have investigated the effect of environmental factors on the spread of COVID-19, most of which focused on temperature. For instance, Briz-Redón and Serranon-Aroca [6] made a comprehensive review of recent reseaches on how climate affects COVID-19's global spread. Their findings revealed that 33 out of 61 articles suggested a negative correlation between COVID-19 and temperature. Few studies reviewed by Briz-Redón and Serranon-Aroca [6] explored the effect of other meteorological factors on COVID-19 such as rainfall, solar radiation and wind speed. One paper out of six found a negative association between solar radiation and COVID-19 transmission; four articles out of eleven suggested a positive association between rainfall and the spread of COVID-19. Endeshaw et al.[7], in their study which examined how the climate factors affect the spread of COVID-19 pandemic in Addis Ababa, Ethiopia, found that factors such as humidity, rainfall, and wind speed have an influence on COVID-19 transmission.

Sera et al.[8] estimated weather-dependent signatures while accounting for socio-economic variables and non-pharmaceutical interventions, in the early stages of the COVID-19 pandemic. In 409 cities across 26 nations, they found a weak non-linear association between mean temperature and the effective reproduction number. They indicated that early interventions have a bigger impact on the reproduction number with a decrease of 0.285 They did mention a little amount of evidence suggesting that local epidemics may have been affected by meteorological conditions in their early stages, but they came to the conclusion that population behavior and government initiatives are the main factors influencing transmission.

Fontal et al.[9] showed, for COVID-19 cases, that there are strong consistent negative impacts of both temperature and absolute humidity at broad regional scales using a statistical method designed to detect transitory associations. They found evidence of strong disease responses in the first two waves, indicating distinct temperature and absolute humidity ranges that are comparable to those previously mentioned for seasonal influenza. For COVID-19, a process-based model that includes a temperature-dependent transmission rate outperforms baseline formulations without a driver or sinusoidal seasonality in all examined locations and pandemic waves. They identified COVID-19 as a seasonal low-temperature disease and assert that the airborne pathway had a significant role in the spread of SARS-CoV-2, which has implications for preventative strategies.

Various studies have utilized various mathematical and statistical models to predict the transmission and intervention consequences. Some researchers have used mathematical models such as deterministic and stochastic models to investigate the dynamics, prediction, prevention, and impact of control measures on the COVID-19 pandemic [10, 11]. Moreover, time series and regression models were also used for predicting and monitoring COVID-19 [12, 13]. For instance, by using the long-term time-dependent epidemiological models SIRD and SEIRD, Manik and Signh [25] explored how temperature and humidity affect the transmission of the virus in several Indian states. In order to determine whether there is any relationship between the effective reproduction number and the temperature, relative humidity, and absolute humidity, they utilized a linear regression approach. In most Indian states, the effective reproduction number has a statistically significant negative connection with both relative and absolute humidity according to their findings.

Data analysis in many different disciplines, including the social, behavioral, and health sciences is associated with the use of General Linear Model (GLM). There are a number of extensions to the GLM, but two main extensions, that are particularly useful in developing a more flexible statistical framework for fixed-effects regression modeling, are the Vector Generalized Additive Models (VGAMs) and the Vector Generalized Linear Models (VGLMs). The VGLM and VGAM framework maintains the benefits of GLMs

 http://dx.doi.org/10.4314/rjeste.v7i1.8

while also smoothing to a much greater extent.In contrast to the long-held belief that common and routine regression modeling in applied statistics has been hindered by a lack of a common framework, the VGAM now includes over 150 family functions, allowing users to freely change model components within a large, flexible framework[5]. Generalized additive models, which are a non-parametric extension of GLMs provide a powerful class of models for data-driven exploratory data analysis [14]. In particular, the classes of VGLMs and VGAMs offer adaptive smoothing within a unified framework, which has significant benefits for extreme value data analysis. Using this class of models, all extreme value distribution parameters can be modeled as smooth or linear functions of covariates [15–17]. More details about VGLMs and VGAMs can be found in [18].

The goal of this article is to highlight the benefit gained from taking into account the classes VGAM and Generalized Extreme Value models (GEV) in the context of extreme value analysis. The aim of this study is to determine and analyse the distribution of extreme weekly cases of COVID-19 in Rwanda during the period of $14^{th}$ March 2020 to $1^{st}$ September 2021 and analyze the effect of climate factors on extreme daily cases of COVID-19. The climate factors considered in this paper are weekly average temperature, weekly average solar radiation, and weekly total rainfall from several locations in Rwanda.

Apart from the introductive section that provides the backgroud of the study , this paper is oganized as follows. The methodology is described in Section 2 and comprises the data description, the statistical method of analysis, and the method of parameter estimation. The presentation and interpretation of the findings for both the exploratory analysis and the use of the Generalized Extreme Value and Vector Generalized Additive Models are covered in the third section. The results are discussed in Section 4, followed by a summary conclusion in Section 5, and acknowledgment in Section 6.

# 2 Materials and Methods

## 2.1 Description of data

This study analysed the weekly maximum new cases together with the climate data as covariates. These are secondary data obtained from Rwanda Biomedical Center (RBC) and Rwanda Meteorological Agency respectively. The raw data for the covariates were recorded as daily maximum and minimum temperature, total daily rainfall and daily solar radiation for different stations across the country while the response variable data were recorded as daily new infections cases from March 2020 to September 2021. This study considered temperature, rainfall and solar radiation as the only covariates based on fact that they were proven to have effect on predicting the transmission of COVID-19 as reported by some authors [6, 7]. The obtained raw data were cleaned and transformed for further analysis. The covariates have been transformed into weekly average temperature (measured in degrees Celcius), weekly total rainfall (measured in millimeters) and weekly average solar radiation (measured in Watt per square metter), respectively whereas the response variable considered is the weekly maximum new cases. Therefore, the sample size used is equal to 78 observations which amounts to the number of weeks spanning the time period considered in this study. The weekly maximum new cases constitute the response variable and the three covariates are the predictors in the model.

## 2.2 Statistical methods for analysis

This section reviews a few details about Vector Generalized Linear Models (VGLMs) and Vector Generalized Additive Models (VGAMs).

### 2.2.1 The extreme value theory and Generalized Extreme Value (GEV) distribution

Let $y_1, y_2, ..., y_n$, be i.i.d. random sample of size $n$ drawn from a distribution $F$, and let $M_n = \max(y_1, y_2, ..., y_n)$. According to Coles [20], the extreme value theory specifies that as $n \to \infty$ the class of non-degenerate limiting distributions for $M_n$, under linear normalization, is the GEV distribution

$$G(z) = \exp\left\{ -\left[1 + \zeta \left(\frac{z-\mu}{\sigma}\right)\right]^{-1/\zeta} \right\}. \tag{1}$$

There are three parameters in equation (1), namely

- location parameter, $-\infty < \mu < \infty$, which is the center of the GEV distribution.
- scale parameter, $\sigma > 0$, which determines the size of deviations of $\mu$, and
- shape parameter, $-\infty < \zeta < \infty$, shows how rapidly the upper tail decays.

$G(z)$ is defined on set $\left\{z : 1 + \zeta\left(\frac{z-\mu}{\sigma}\right)\right\}$.

In this formula, positive $\zeta$ implies a heavy tail while negative one implies a bounded tail and the limit of $\zeta \to 0$ implies an exponential tail [20].

### 2.2.2 Linear and additive models

Yee and Stephenson [19] defined the Vector Generalized Linear Models (VGLMs) as a model for which the conditional distribution of $\mathbf{y}$ given exploratory (covariates) $\boldsymbol{x}$ is of the form

$$f(\boldsymbol{y}|\boldsymbol{x}; \boldsymbol{B}) = h(\boldsymbol{y}, \eta_1, \ldots, \eta_M),$$

 http://dx.doi.org/10.4314/rjeste.v7i1.8

for some known function $h(.)$, where $\boldsymbol{B} = (\boldsymbol{\beta_1}, \ldots, \boldsymbol{\beta_M})$ is a $p \times M$ matrix of unknown regression coefficients, and the $j^{th}$ linear predictor is

$$\eta_j(\boldsymbol{x}) = \boldsymbol{\beta}_j^T \boldsymbol{x} = \sum_{k=1}^{p} \beta_{(j)k} x_k, j = 1, \ldots, M, \tag{2}$$

where $\boldsymbol{x} = (x_1, \ldots, x_p)^T$, $\boldsymbol{y}$ is an observed response vector, $M$ is the number of parameters to be estimated in the extreme value models, and $p$ is the number of covariates.

For simple extreme value model like stationary GEV and General Pareto (GP), $M = 3$ or $M = 2$ for example. The VGLMs are thus like GLMs but allow for multiple linear predictors, and they encompass models outside the limited confines of the exponential family. Yee and Stephenson [19] proposed that the $\eta_j$ of VGLMs may be applied directly to parameters of a distribution rather than just to means as for GLMs. A simple example is a univariate distribution with a location parameter $\mu$ and a scale parameter $\sigma > 0$, where we may take $\eta_1 = \mu$ and $\eta_2 = \log \sigma$. Yee, Yee and Stephenson [18, 19] provided that the VGAMs are the additive-model extensions to VGLMs, that is, equation (2) is generalized to

$$\eta_j(\mathbf{x}) = \beta_{(j)k} + \sum_{k=1}^{p} f_{(j)k}(x_k), \tag{3}$$

a sum of smooth non-linear functions $f_{(j)k}(x_k)$ of the individual covariates.

Simon et al. [21] showed that the model smoothness $f_{(j)k}(x_k)$ is transformed into a linear model via basis expansions of modest rank $(D_k)$ as follows

$$f_{(j)k}(x_k) = \sum_{d=1}^{D_k} \beta_{kd} b_{kd}(\mathbf{x}), \tag{4}$$

where $\beta_{kd}$ represent the unknown coefficients and $b_{kd}$ represents the known basis functions such as splines, usually chosen to have good approximation theoretic properties. Substitute equation (4) in equation (3), we get

$$\eta_*(\mathbf{x}) = \beta_0 + \sum_{k=1}^{p} \sum_{d=1}^{D_k} \beta_{kd} b_{kd}(\mathbf{x}).$$

The basis representations can be written in matrix notations as follows

$$\eta_* = \mathbf{x}^T \mathbf{B},$$

where $\mathbf{x}^T$ is a row of a design matrix $\mathbf{X}$ which has elements determined by the choice of $b_{kd}$ basis functions and $1 + \sum_{k=1}^{p} D_k$ columns, each of which corresponds to an element of

**B**. For the GEV, $\mu(x) = \eta_\mu(\mathbf{x}), \log\sigma(\mathbf{x}) = \eta_\sigma(\mathbf{x})$ and $\zeta(\mathbf{x}) = \eta_\zeta(\mathbf{x})$, see [19].

### 2.2.3 Parameters estimation

This subsection describes the method of estimation of parameter for GEV-VGAM model. Consider a model for an $n-$ vector of observations, $\boldsymbol{y}$, constructed in terms of unknown parameters $\boldsymbol{\theta}(\boldsymbol{x}) = (\mu(\boldsymbol{x})), \sigma(\boldsymbol{x}), \eta(\boldsymbol{x})$, and some unknown function $\eta_j$ defined in equation (3), of covariates, $x_j$. Youngman [22] proposed that the estimation of GEV corresponds to the estimation of basis coefficients. That is,the likelihood function for model (3) can be denoted by $L(\eta_j(\boldsymbol{x}), \boldsymbol{Y}) = L(\boldsymbol{B})$ with log-likelihood $l(\boldsymbol{B})$.

Authors in [19, 21, 22] and [23] proposed the penalty log-likelihood function as follows

$$\mathcal{L}(\boldsymbol{B}_{\boldsymbol{\lambda}}, \boldsymbol{\lambda}) = l(\boldsymbol{B}_{\boldsymbol{\lambda}}) - \frac{1}{2}\boldsymbol{B}'\boldsymbol{S}_{\boldsymbol{\lambda}}\boldsymbol{B}, \tag{5}$$

where $\boldsymbol{\lambda} = (\lambda_1, \lambda_2, \ldots, \lambda_k)$ is a smoothing parameter, $\boldsymbol{S}_{\boldsymbol{\lambda}}$ is a penalty matrix with elements determined by the chosen $b_{kd}$ basis functions with $\boldsymbol{S}_{\boldsymbol{\lambda}} = \sum_{k=1}^{K}\lambda_k\boldsymbol{S}_k$.

The estimation of parameter can be estimated by maximizing the penalty log-likelihood function defines in equation (5). That is,

$$\widehat{(B)} = argmax_{(B)}\mathcal{L}(\boldsymbol{B}_{\boldsymbol{\lambda}}, \boldsymbol{\lambda}).$$

The GEV-VGAM model of the research problem was formulated using the data described in section (2.1), where the response variable is the *weekly infected new cases* and the three considered covariates are *total weekly rainfall*, *average weekly solar radiation* and *average weekly temperature*. In practice we may wish to constrain the effect of a co-variate to be the same for some of the $\eta_j$ with $j = 1, 2, 3$ and to have no effect for others. For example, for VGAMs we may wish to take

$$\begin{aligned}
\eta_1 &= \beta_{10} + f_{(1)1}(x_1) + f_{(1)2}(x_2) + f_{(1)3}(x_3);\\
\eta_2 &= \beta_{20} + f_{(2)1}(x_1) + f_{(2)2}(x_2) + f_{(2)3}(x_3);\\
\eta_3 &= \beta_{30} + f_{(3)1}(x_1) + f_{(3)2}(x_2) + f_{(3)3}(x_3).
\end{aligned}$$

We will relate the $\eta_j, j = 1, 2, 3$ with the model parameters after fitting the GEV-VGAM models in the next section.

# 3 Results

This section describes the exploratory data analysis of the response variable with their covariates and the Generalized Extreme Value and Vector Generalized Additive Model (GEV-VGAM). The exploratory data analysis deals with the descriptive statistics of

 http://dx.doi.org/10.4314/rjeste.v7i1.8

weekly maximum infected new cases and covariates, time series plot of the weekly maximum infected new cases versus weeks, and the scatter plots of the weekly maximum infected new cases versus the covariates. The estimation of parameters and the fitted model are reviewed in GEV-VGAM sub-section.

## 3.1 Exploratory data analysis

The descriptive statistics of weekly maximum infected new cases of COVID-19 with their climate variables and time series plot of the weekly maximum infected new cases and the climate variables versus weeks are presented in Table 1 and Figure 1 respectively. The used data are collected from the week of 8-14/03/2020 to the week of 1-4/9/2021 (week 78).

Table 1: Weekly maximum new cases of COVID-19 and climate variables

| **Variables** | **Minimum** | **Maximum** | **Mean** | **SD** |
|---|---|---|---|---|
| Weekly maximum new cases (number) | 1 | 3072 | 257 | 531.47 |
| Weekly average temperature ($^oC$) | 18.66 | 21.6 | 19.95 | 0.59 |
| Weekly average solar radiation (W/$m^2$) | 59.81 | 89.11 | 76.32 | 7.26 |
| Weekly total rainfall (mm) | 0.001 | 13.6 | 3.08 | 3.18 |

* SD: Standard Deviation

Within this period, the total number of 20722 cases of weekly maximum number of infections of COVID-19 were recorded in this study in which the mean of weekly maximum number of confirmed cases was 257 (maximum =3072), during which the weekly average temperature ranged from $18.66^0C$ to $21.6^0C$ with the overall mean equal to $19.95^0C$. The overall mean of weekly average solar radiation and weekly total rainfall were $3.08mm$ (maximum = $13.6mm$) and $76.32W/m^2$ (maximum=$89.11W/m^2$), respectively.

Figure 1 shows the time series plot of weekly maximum number of confirmed cases of COVID-19, weekly average temperature, weekly total rainfall and weekly average solar radiation. It is seen that the weekly maximum number of confirmed cases of COVID-19 increased slowly in the first twenty weeks. This increment might have due to the measures adopted by the Government of Rwanda like wearing mask, physical social distancing, public lock-down, quarantine, etc. This Figure also reveals that from week 67 (13-19/06/2021) up to week 74 (8-14/08/2021), the weekly maximum number of confirmed cases of COVID-19 increased rapidly due to the new variant of COVID-19 called Delta. This variant spread rapidly and it is more contagious than the other variants. The weekly maximum of confirmed COVID-19 infections dramatically decreased after week 74 as a result of the Rwandan government's decision to vaccinate every Rwandan starting with

 http://dx.doi.org/10.4314/rjeste.v7i1.8

the eldest. In addition, Figure 1 shows that the peaks in the weekly maximum number of confirmed cases of COVID-19 against the climate variables are correlated.

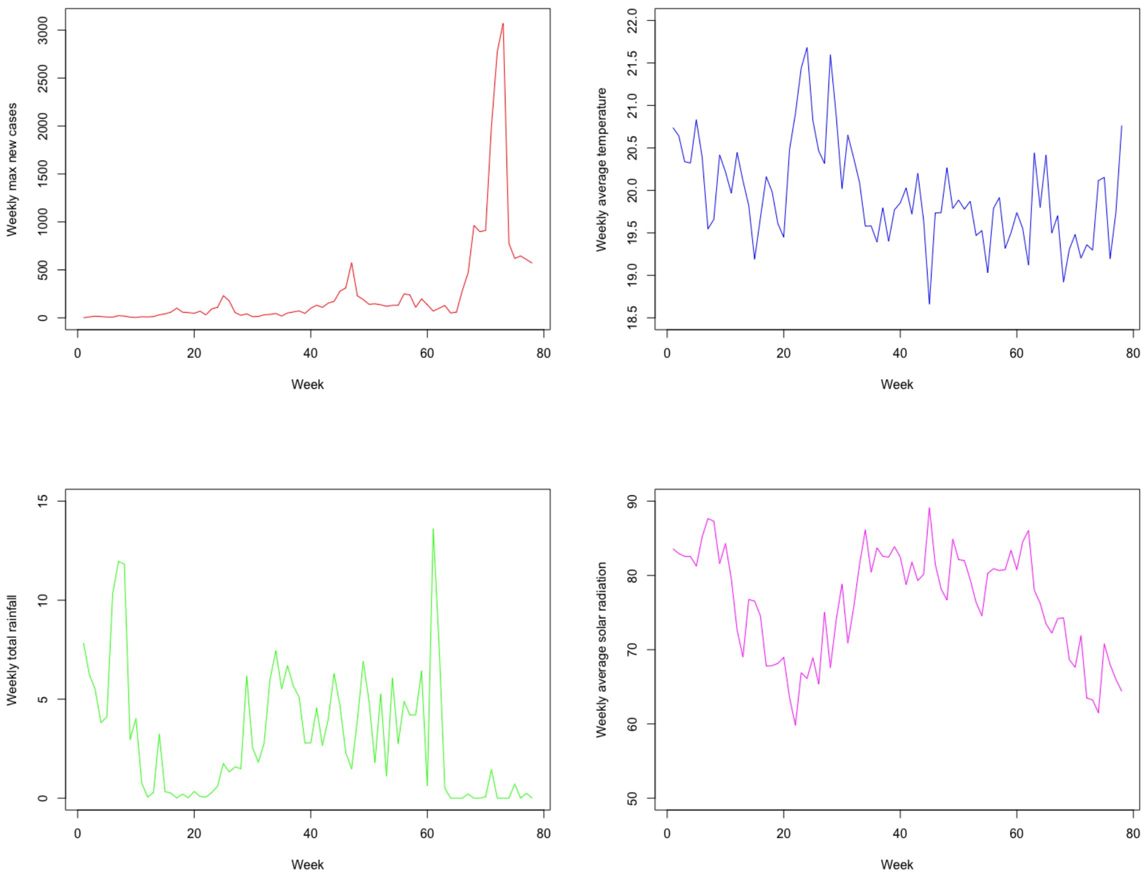


Figure 1: Time series plots of weekly maximum new cases and climate factors

The following figures represent the scatter plots of weekly maximum new cases of COVID-19 against climate variables.

Figure 2 shows that a higher temperature is likely to reduce the number of new infection cases of COVID-19 and a lower solar radiation increases the number of new infection cases of COVID-19, see Figure 4. Figure 3 reveals that an increase in rainfall may reduce the weekly maximum new cases of COVID-19 due to the case that too much rain could generate incentives to stay indoors. Therefore, a negative relationship between weekly maximum new cases of COVID-19 and our climate variables was identified.

 http://dx.doi.org/10.4314/rjeste.v7i1.8

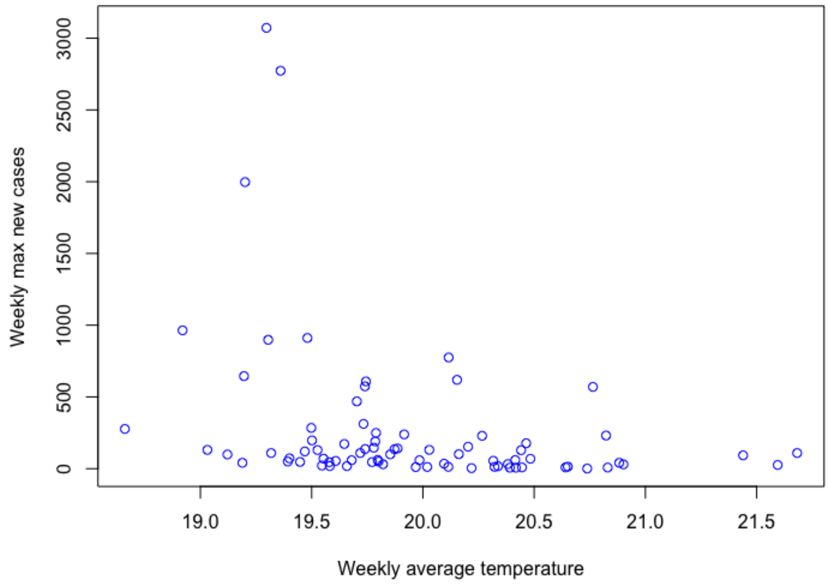


Figure 2: Scatter plot of weekly maximum new cases against weekly average temperature($^oC$)

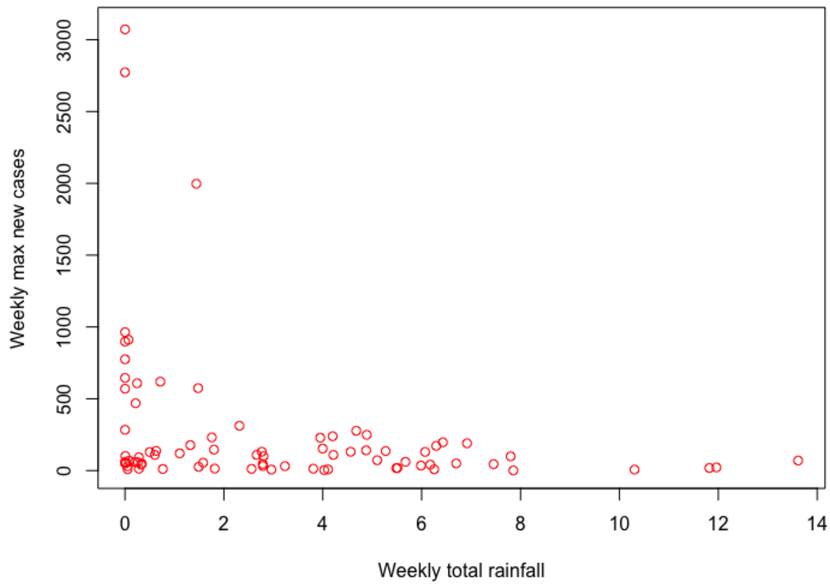


Figure 3: Scatter plot of weekly maximum new cases against weekly total rainfall($mm$)

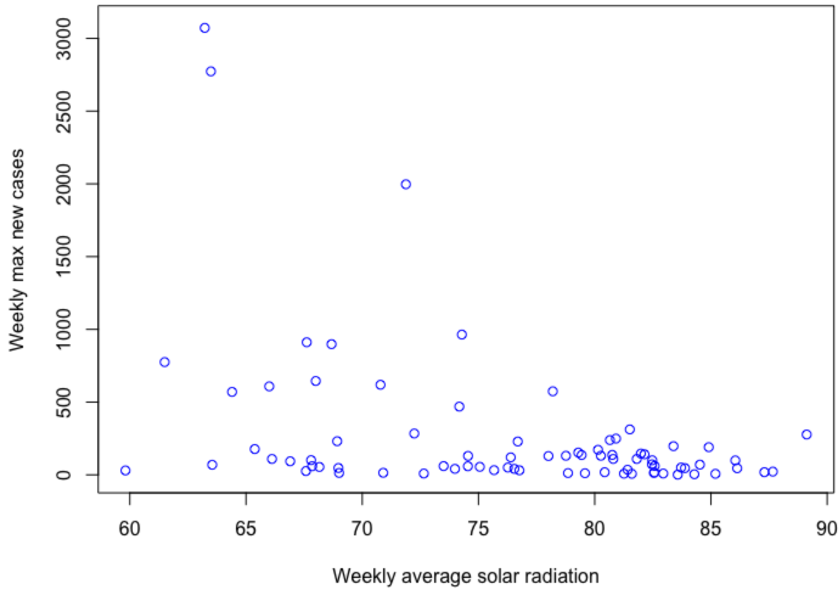


Figure 4: Scatter plot of weekly maximum new cases against weekly average solar radiation($W/m^2$)

## 3.2 Generalized Extreme Value and Vector Generalized Additive Model (GEV-VGAM)

In this subsection, the GEV-VGAM fitted with COVID-19 and climate data is described.

### 3.2.1 GEV-VGAM parameter estimates

The GEV-VGAM was fitted to weekly maximum new infections and covariates. The results are presented in Table (2).

 http://dx.doi.org/10.4314/rjeste.v7i1.8

*Rwanda Journal of Engineering, Science, Technology and Environment, Volume 7, Issue1, March 2025*
*eISSN: 2617-233X |print ISSN: 2617-2321*

Table 2: Results of the fitted GEV-VGAM for aall considered climate factors

| Parameter | Estimate | Stand.Error | $t$-value | P-value |
|---|---|---|---|---|
| **Parametric terms** | | | | |
| location | 97.39 | 0 | 60631.54 | <2e-16 |
| logscale | 4.7 | 0.06 | 80.86 | <2e-16 |
| shape | 0.56 | 0.07 | 0 7.9 | <1.41e-15 |
| **Smooth terms** | | | | |
| *Location* | | | | |
| Smooth spline | Effective df | Max.df | Chi.sq | P-value |
| s(Temperature) | 0 | 9 | 53.79 | <2.23e-13 |
| s(Rainfall) | 0 | 9 | 5.72 | 0.0168 |
| s(Solar-radiation) | 0 | 9 | 6.83 | 0.00894 |
| *logscale* | | | | |
| s(Temperature) | 2.63 | 9 | 11.15 | 0.0418 |
| s(Rainfall) | 2.85 | 9 | 18.98 | 0.000148 |
| s(Solar-radiation) | 2.04 | 9 | 0.98 | 0.508 |
| *shape* | | | | |
| s(Temperature) | 0.99 | 9 | 0.54 | 0.462 |
| s(Rainfall) | 0.83 | 9 | 16.58 | <4.66e-05 |
| s(Solar-radiation) | 1.05 | 9 | 0.11 | 0.601 |


http://dx.doi.org/10.4314/rjeste.v7i1.8

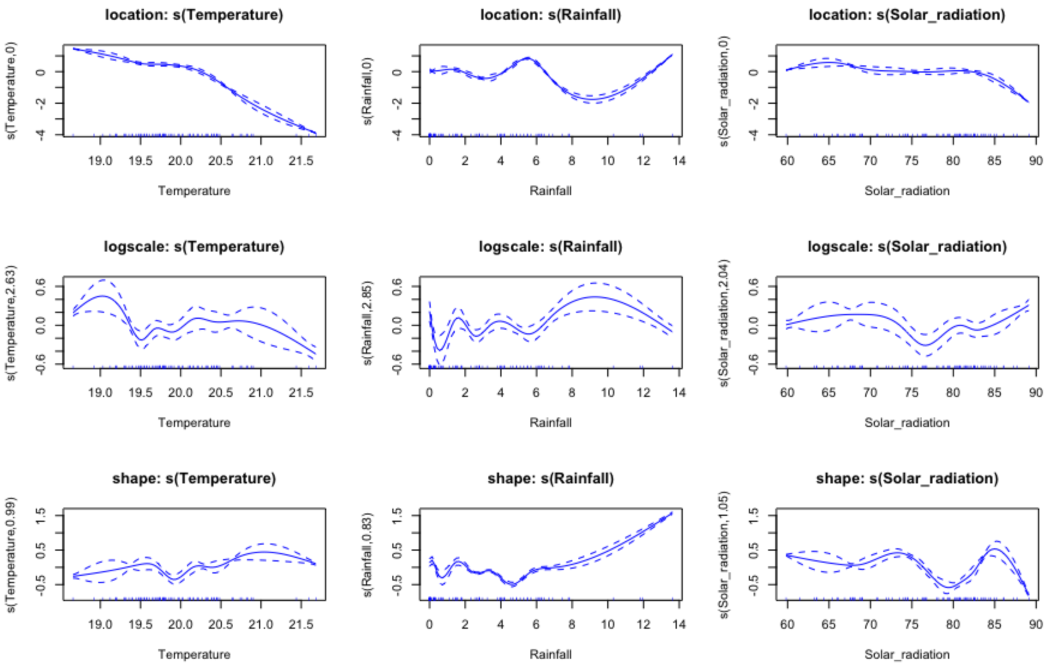


Figure 5: **GEV-VGAM fitted to weekly maximum new cases for all climate factors**

The results show that the effective degrees of freedom (edf) for location are zero, which means that the location of the GEV distribution is not predicted by the considered covariates. For the scale, the edf show high non-linear relationship ($edf > 2$ ) with all three covariates, and the smoothing spline terms are significant except for solar radiation (p-value= 0.508). For the shape, the edf are close to one indicating a linear relationship. However, the effects of both temperature and solar radiation on the shape of the GEV are not significant. These results suggest a GEV model with a constant location parameter, a linear model of the shape with rainfall as a predictor, and a non-linear model of the scale with temperature and rainfall as predictors is more appropriate and its results is presented in Table (3) and the fitted smooth is presented in Figure (6)


http://dx.doi.org/10.4314/rjeste.v7i1.8

Table 3: Results of the fitted GEV-VGAM for significant climate factors

**Parametric terms**

| Parameter | Estimate | Stand.Error | $t$-value | P-value |
|---|---|---|---|---|
| location | | | | |
| (intercept) | 53.26 | 0.11 | 478.37 | <2e-16 |
| logscale | | | | |
| (intercept) | 5.23 | 0.04 | 119.34 | <2e-16 |
| shape | | | | |
| (intercept) | 6.07 | 0.07 | 89.02 | <1.41e-15 |
| (intercept) | 0.07 | 0.01 | 10.21 | <1.41e-15 |

**Smooth terms**

*Location*

| Smooth spline | Effective df | Max.df | Chi.sq | P-value |
|---|---|---|---|---|
| s(Temperature) | 8.95 | 9 | 397030490 | <2e-13 |
| s(Rainfall) | 9 | 9 | 13187619 | <2e-13 |

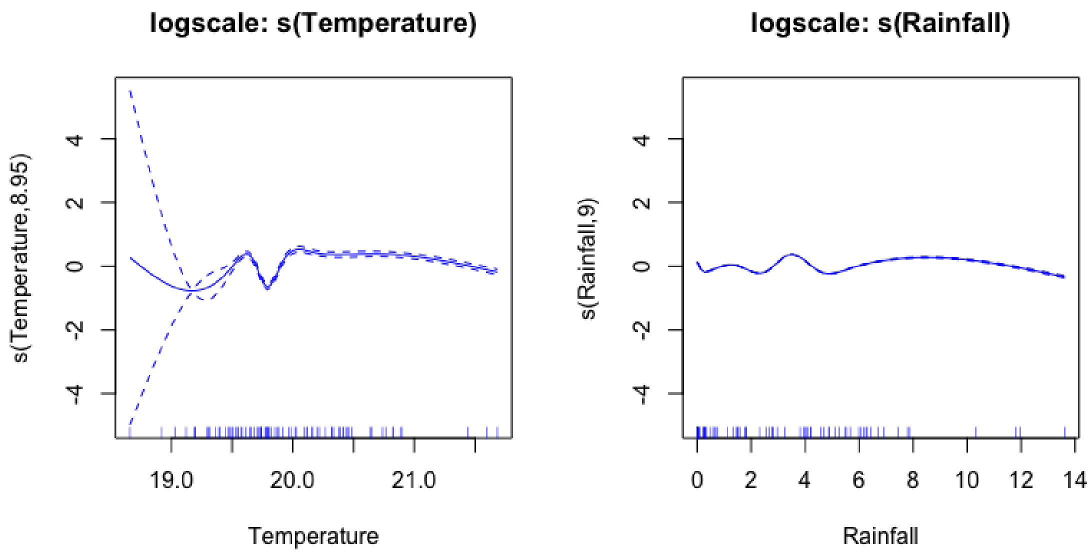


Figure 6: **GEV-VGAM fitted to weekly maximum new cases for significant climate factors**


http://dx.doi.org/10.4314/rjeste.v7i1.8

# 4 Discussion

The findings indicate that an increase in solar radiation and the high temperature may decrease the weekly maximum number of new COVID-19 infections. Some previous studies have looked at the impact of climate environment factors on the spread of COVID-19. For example Patwary et al. [24] looked into the effects of socio-demographic and environmental factors on COVID-19 transmission. The results of rainfall varied between investigations. Population density, educational attainment, and income were identified as potential contributors to the coronavirus outbreak among the non-climatic variables. A substantial link between climatic and sociodemographic parameters and the COVID-19 outbreak was observed. In similar studies, Endeshaw et al.[7] found that the number of COVID-19 cases showed seasonal variation, with the rainy season seeing the highest number of cases reported and the dry season reporting the lowest. Using articles and data that analyzed and investigated the climatic and environmental factors of COVID-19 in African countries, Mwiinde et al. [26] made a systematic review that explored the climatic and environmental factors influencing COVID-19 transmission in Africa. The results showed that there is evidence suggesting the influence of climatic and environmental factors on the spread of COVID-19 in Africa; however, the evidence requires further investigation in all six regions of Africa and at the country level to comprehend the role of weather patterns and environmental factors in the transmission of COVID-19.

The results show that GEV-VGAM model with climate factors is reliable to predict extreme values of new infections of COVID-19 in Rwanda. The findings show a linear relationship of the shape and nonlinear relationship of the scale with climate factors, while the location parameter is constant. However, there are non significant effects of both temperature and solar radiation on the shape parameter of the GEV, which means that an increase on temperature and solar radiation does not imply an increase of new infections of COVID-19. In similar studies, Sezer et al. [27] used generalized additive models for location, scale, and form factors to explain the nonlinear relationship between the daily quantity of intense rainfall and relevant predictor variables. The model's approximated mean function converges to the real mean function, according to their results. In their study, Utami et al. [28] used the VGAM to examine data on extreme rainfall in Indramayu, Indonesia (VGAM). The findings showed that the VGAM with General Pareto Distribution is capable of reliably forecasting extreme rainfall data.

# 5 Conclusion

Since the COVID-19 pandemic first emerged in 2020, there have been concerns regarding how to control it and prevent it from spreading. In order to guide health care providers, policymakers, and governing bodies for interventions, it is crucial to examine the dynamics of COVID-19 and identify various factors, especially those linked to extreme positive

 http://dx.doi.org/10.4314/rjeste.v7i1.8

cases. The objective of this paper was to examine the distribution of COVID-19 extreme weekly cases in Rwanda for the period of 14-th March 2020 to 1st September 2021 as well as the impact of climate conditions on COVID-19 extreme daily infection cases. This study analyzed weekly maximum positive COVID-19 cases together with three climate covariates (temperature, rainfall, and solar radiation) to examine the predictive effect of climate factors on extreme COVID-19 cases. The findings showed that using climate factors as covariates, Generalized Additive Extreme Value Models are useful for predicting determinant parameters of COVID-19 transmission in Rwanda such as daily infection cases. According to the results, a Generalized Extreme Value distribution with a constant location parameter, a linear model for the shape parameter with rainfall as a predictor, and a non-linear model for the scale parameter with temperature and rainfall as predictors fit the weekly maximum positive cases the best. Thus, the dynamics of the COVID-19 pandemic are significantly influenced by the weather, particularly the temperature and rainfall. However, according to Schuster Arthur [29], periodicity and volatility are the inherent features of weather factors, particularly temperature, hence the present results should not be viewed as a generalization. The results from this research may be used to enhance COVID-19 spread prediction and plan for interventions in Rwanda, specifically for extreme cases of infection. In future studies, one can analyse the effect of climate factors on extreme COVID-19 death, the effect of control measures and vaccination on the spread of COVID-19. Future research might also focus on how human behavior influences the transmission of COVID-19, as well as how travel habits and governmental control measures influence the spread of this pandemic. Using autocorrelation concepts like the Hurst exponent or the Entropy, one can also compare the roughness of epidemiological data with that of meteorological data.

## 6 Acknowledgment

The authors acknowledge the National Council for Science and Technology for funding this research; Rwanda Biomedical Centre and Meteo Rwanda for availing the used data.

# References

[1] Mingwang Shen, Zhihang Peng, Yanni Xiao, and Lei Zhang. Modeling the epidemic trend of the 2019 novel coronavirus outbreak in china. *The Innovation* **2020**, **1(3)**.

[2] Güner, Hatice Rahmet and Hasanoğlu, İmran and Aktaş, Firdevs. Covid-19: Prevention and control measures in community. *Turkish Journal of medical sciences* **2020**, **50(9)**, 571–577.

[3] Taeyong Lee, Hee-Dae Kwon, and Jeehyun Lee. The effect of control measures on covid-19 transmission in south korea. *PloS one*, **2021**, **16(3)**:e0249262.

[4] Brian Fleischer, Ronald Olum, Frederick Nelson Nakwagala, Dianah Rhodah Nassozi, Ivaan Pitua, Elijah Paintsil, Joseph Baruch Baluku, and Felix Bongomin. Higher intensive care unit consultations for covid-19 patients living with hiv compared to those without hiv co-infection in uganda. *Journal of Medical Virology*, **2022**.

[5] Yee, T. W. (2015). *Vector generalized linear and additive models: with an implementation in R* (Vol. 10, pp. 978-1). New York: springer.

[6] Briz-Redón, Álvaro and Serrano-Aroca, Ángel. The effect of climate on the spread COVID-19 pandemic: A review of findings, and statistical and modeling techniques. In *Progress in Physical Geography: Earth and Environment*, **2020**, **44(5)**, 591–604.

[7] Endeshaw, F. B., Getnet, F., Temesgen, A. M., Mirkuzie, A. H., Olana, L. T., Alene, K. A., & Birhanie, S. K. Effects of climatic factors on COVID-19 transmission in Ethiopia.*Scientific Reports*, **2022**, **12(1)**, 1-10.

[8] Sera, F., Armstrong, B., Abbott, S., Meakin, S., O'Reilly, K., von Borries, R., ... & Lowe, R. (2021). A cross-sectional analysis of meteorological factors and SARS-CoV-2 transmission in 409 cities across 26 countries. *Nature communications*, **12(1)**, 5968.

[9] Fontal, A., Bouma, M. J., San-José, A., López, L., Pascual, M., & Rodó, X. (2021). Climatic signatures in the different COVID-19 pandemic waves across both hemispheres. *Nature Computational Science*, **1(10)**, 655-665.

[10] Wenyu Zhao, Yongjian Zhu, Jingui Xie, Zhichao Zheng, Haidong Luo, and Oon Cheong Ooi.The moderating effect of solar radiation on the association between human mo- bility and covid-19 infection in europe. *Environmental Science and Pollution Research* **2022**, **29(1)**, 828–835.

[11] Shah Hussain, Elissa Nadia Madi, Hasib Khan, Haseena Gulzar, Sina Etemad, Shahram Rezapour, and Mohammed KA Kaabar.On the stochastic modeling of covid-19 under the environmental white noise. *Journal of Function Spaces*, **2022**.

[12] Xiaofeng Liu, Zubair Ahmad, Ahmed M Gemeay, Alanazi Talal Abdulrahman, EH Hafez, and N Khalil. Modeling the survival times of the covid-19 patients with a new statistical model: A case study from china. *PloS one*, **2021**, **16(7)**:e0254999.

[13] Anh Ngoc Nguyen, Xuan Thi Thanh Le, Nhung Thi Kim Ta, Danny Wong, Nguyen Thao Thi Nguyen, Huong Thi Le, Thao Thanh Nguyen, Quan Thi Pham, Quynh

Thi Nguyen, Quan Van Duong, et al. Knowledge and self-protective practices against covid-19 among healthcare workers in vietnam. *Frontiers in Public Health*, **2021**, **9**.

[14] Marc Schneble, Giacomo De Nicola, Goöran Kauermann, and Ursula Berger. A statistical model for the dynamics of covid-19 infections and their case detection ratio in 2020. *Biometrical Journal*, **2021**, **62(8)**, 1623–1632.

[15] Richard L Smith. Extreme value theory based on the r largest annual events. *Journal of Hydrology*, **1986**, **86(1-2)**, 27–43.

[16] Ishay Weissman. Estimation of parameters and large quantiles based on the k largest observations. *ournal of the American Statistical Association*, **1978**, **73(364)**, 812–815.

[17] Jonathan A Tawn. Bivariate extreme value theory: models and estimation. *Biometrika*, **1988**, **75(3)**, 397–415.

[18] Thomas W Yee. *Vector generalized linear and additive models: with an implementation in R*; volume 10; Springer: New York, USA, 2015; pp.978-1.

[19] Thomas W Yee and Alec G Stephenson. Vector generalized linear and additive extreme value models. *Extremes*, **2007**, **10(1)**, 1–19.

[20] Stuart Coles, Joanna Bawa, Lesley Trenner, and Pat Dorazio. *An introduction to statistical modeling of extreme values*; volume 10208; Springer: London, UK, 2001; p.208.

[21] Simon N Wood, Natalya Pya, and Benjamin Säfken. Smoothing parameter and model selection for general smooth models. *Journal of the American Statistical Association*, **2016**, **111(516)**, 1548–1563.

[22] Benjamin D Youngman. Evgam: An r package for generalized additive extreme value models. *ArXiv preprint arXiv:2003.04067*, **2020**.

[23] Mari R Jones, Stephen Blenkinsop, Hayley J Fowler, David B Stephenson, and Christo- pher G Kilsby. Generalized additive modelling of daily precipitation extremes and their climatic drivers. *National Center for Atmospheric Research*, Colorado, **2013**.

[24] Muhammad Mainuddin Patwary, Mondira Bardhan, Asma Safia Disha, Mehedi Hasan, Md Zahidul Haque, Rabeya Sultana, Md Riad Hossain, Matthew HEM Browning, Md Ashraful Alam, and Malik Sallam. Determinants of covid-19 vaccine acceptance among the adult population of bangladesh using the health belief model and the theory of planned behavior model. *Vaccines*, **2021**, **9(12)**,1393.

[25] Moumita Manik and Rakesh K Singh. Role of toll-like receptors in modulation of cytokine storm signaling in sars-cov-2-induced covid-19. *Journal of medical virology*, **2022**, **94(3)**, 869–877.

[26] Mwiinde, A. M., Siankwilimba, E., Sakala, M., Banda, F., & Michelo, C. (2022). Climatic and Environmental Factors Influencing COVID-19 Transmission—An African Perspective. *Tropical Medicine and Infectious Disease*, **7(12)**, 433.

[27] Sezer, Ahmet and Kan Kılınç, Betül and others. Modelling extreme rainfalls using generalized additive models for location, scale and shape parameters. *Applied Ecology and Environmental Research*, **2016**, **14(4)**, 635–644.

[28] Eka Putri Nur Utami, Aji Hamim Wigena, and Anik Djuraidah. Vector generalized additive models for extreme rainfall data analysis (study case rainfall data in indramayu). *AIP Conference Proceedings*, **2016**, volume 1707, no. 1, p. 080007. AIP Publishing LLC.

[29] Schuster A. On the investigation of hidden periodicities with application to a supposed 26 day period of meteorological phenomena. *Journal of Geophysical Research*, **3(1)**, 13–41.