Evaluation of demographic and climatic conditions in the global spread of coronavirus disease 2019 (COVID-19)

Prabhat Kumar Trivedi[1], Ranadeep Datta[1], Ayush Kumawat[1], Rahul Kumar[1], Isha Bhardwaj[1], Varun Agiwal[1], Saurabh Kumar[1], Ashok Kumar[1], Ashutosh Shukla[2,] Jayendra Nath Shukla[1,a] and Jitendra Kumar[1a],

[1]*Central University of Rajasthan, Bandarsindri, Kishangarh, Ajmer, Rajasthan, India*

[2]*National Statistical Office, Field and Operationn Division, Agra, Uttar Pradesh, India*

ABSTRACT

There has been a history of outbreak of viruses from time to time like emergence of Ebola virus, H1N1, Astrovirus etc. Coronaviruses (CoVs) having RNA as their genetic material came into limelight during the outbreak of Severe Acute Respiratory Syndrome-CoV (SARS-CoV) in 2002 and further during the outbreak of Middle East respiratory syndrome coronavirus (MERS-CoV) in 2012. Recent outbreak of 2019 novel coronavirus (SARS-CoV-2) in Wuhan, China and further its spread throughout the world has become a global threat to the existence of mankind. Different reports have suggested the relatively rapid spread of SARS-CoV-2 (novel coronavirus) and the disease caused by it, the coronavirus disease 2019 (COVID-19). This leads to the official declaration of COVID-19 by World Health Organization (WHO) as a pandemic disease. Information on the spread of COVID-19 in a particular geographical area and comparison of its spread with respect to the demographic and climatic conditions in different geographical regions is the need of the time to device better strategies to combat the novel coronavirus and also during the emergence of new viral strains in future. Our in depth analysis suggests that the number of COVID-19 cases follows a log normal distribution in each temperature intervals of both minimum and maximum temperature interval except highest interval of minimum temperature interval.

[a]Corresponding Author

Email:

Jayendra Nath Shukla – jayendrashukla@curaj.ac.in

Jitendra Kumar - jitendravkarma@curaj.ac.in

1. INTRODUCTION

The discovery of coronaviruses (CoVs) belonging to Family- *Coronaviridae*, dates back to 1960's (1). CoVs are single stranded RNA viruses (2). The outbreak of CoVs in year 2002 in China caused respiratory infections "severe acute respiratory syndrome (SARS)" (3, 4); this virus was named SARS-CoV. The second outbreak of similar kind of virus (MERS-CoV) resulted in "Middle East respiratory syndrome (MERS)" in the Middle East population (5, 6). A recent outbreak of novel coronavirus (SARS-CoV-2) and the disease spread by it (COVID-19) took place in the Wuhan city of China in December, 2019 which lead to more than 40,000 deaths till March 31, 2020.The causative coronavirus of this outbreak was named SARS-CoV-2 due to its resemblance to SARS-CoV (7–9). SARS-CoV-2 captured the whole world within few months from the start of its emergence. Considering the severity of SARS -CoV-2 and mortality rate of COVID-19 WHO officially declared it to be a pandemic disease.

The source of initial infection of N-CoVs to human population remains elusive. Several mammals (camels, cattle, cats and bats etc) and birds acts as natural host for CoVs (7). Studies suggested the possible spread of NCoVs through bats as they are the natural reservoir of a wide variety of CoVs, including SARS-CoV and MERS- CoV viruses (8–9). The spread of the NCoVs in humans resulted probably because of their direct contact with host animals or consumption of wild animals. However, the actual source(s) and transmission route(s) of SARS-CoV-2 is a matter of further investigation.

Available data suggested the spread of N-CoVs through the community transmission mode, i.e., the spread of the disease when coming in contact with the patients or incubation carriers. Temperature is known to play a very significant role in the community transmission of viral infections (10, 11, 12). This lead us to study the progression of COVID-19 in different geographical locations with respect to environmental temperature of those locations. In the study presented here we analyzed COVID-19 cases (available on 31March 2020) with respect to the geographical area and the human population. We presented the number of total positive cases as on March 31, 2020 in 'per Billion population' and in 'per 10000 Km^2 area'. Further, we modelled the number of cases of various countries based on respective minimum and maximum temperature intervals. As on date, almost all countries are infected and most of the cases are recorded for ambient temperature intervals i.e., maximum temperature interval - 17 to 34 ^oC and minimum temperature interval - 14.5 to 29.5 ^oC. For modeling we explored various distributions like Exponential, Cauchy, Gamma, Lognormal, Logistic and Normal etc. We obtained 'empirical mean and variance' and recorded that in our modelings. We observed that the average number of COVID-19 infections is decreasing with the increase of temperature. Similar trend was also observed for maximum temperature intervals.

2. MATERIAL AND METHODS

2.1 Data Source: The public attention of COVID-19 came into picture on December 31, 2019 when WHO office of China detected the cases of pneumonia with unknown etiology in Wuhan City, (Hubei Province) of China. A total of 44 patients with the similar disease were noticed within a short span of time (four days). This is the time when causal agent of this disease was not identified; WHO declared this as COVID-19 infection on January 20, 2020. The causal agent for COVID-19 was identified as SARS-CoV-2. Fast spread of COVID -19 crossed the boundaries of different geographical locations and infected people in almost all the countries within few months. The data regarding number of COVID-19 patients, human population and climatic conditions of a particular geographical location was collected from various sources listed below:

S. No.	Information	Source

1	Country Code	https://countrycode.org/
2	COVID-19 Data	https://www.worldometers.info/coronavirus/
3	Temperature	https://www.timeanddate.com/
4	Land Area & Population	https://www.worldometers.info/world-population/population-by-country/
5	WHO	WHO Corona virus disease 2019 (COVID-19) Situation Report

'Worldometer' a database provides country wise real-time statistics of various parameters. The details were obtained from https://www.worldometers.info/about/.

2.2 Software used: Since the present study has targeted to explore the COVID-19 cases with respect to different countries, the data preparation is copied as "csv" format and arranged in a country wise manner in Excel sheet. The analysis and fitting of distribution was made using R software and the pictorial presentation was prepared on Excel and Tableau.

In order to correlate the number of COVID-19 cases in different countries with respect to temperature of those countries, the distribution of minimum and maximum temperatures was carried out. Further, the distribution of total COVID-19 cases was obtained considering number of COVID-19 cases as a variable. Same analysis was carried out for the cases in 'per billion population' and in 'per 10000 Km^2 area' to get more generalised view of incidence of COVID-19 with respect to temperature.

3. RESULT AND DISCUSSION

3.1 COVID 19 Worldwide: More than 7.5 Lakh confirmed cases of COVID-19 covering 202 countries, areas or territories was reported till 31 March 2020. This data is inclusive of more than fifty seven thousand new cases on March 31, 2020 itself (WHO situation report 71). Also, more than forty thousand deaths were reported in various countries due to COVID-19 infection. Figure 1 gives the pictorial representation of the intensity of confirmed cases and total deaths in different countries.

Figure-1

We studied the date wise spread and the percentile of number of countries for COVID-19 infections, starting from 31 Dec 2019 to 31 March 2020. Our result shows that only about 10% of total countries were infected with COVID-19 by 30 January, 2020. The infection kept on increasing, covering more number of countries/area/region with time. About 85% of the total countries (i.e.,202 countries) and dependent territories got COVID-19 infection by 31 March, 2020 (Figure 2). Our results also suggest the rapid spread of COVID-19 in the month of March from 30% of the countries to almost all of the countries.

Figure-2

The global human population is estimated to be 7.77 billion. We explored the population and area under the threat of COVID-19 at different dates. It is important to explore this domain as a single infected person in a country is a threat to whole population of that country and land area under threat could be useful to check its spread along territories. Figure 3 represent the population and area wise coverage of COVID-19 cases, worldwide.

Figure-3

It can be observed from 'Figure 3' that initially about 1.5 billion of world's total population was under threat of COVID-19, and this number includes the population of China and Iran. In just 13 days from the date when first coronavirus case was reported (i.e. on 31st of Dec 2019) about 20% of world population (i.e. around 2 billion human population) got the COVID-19 infection after which a sluggish growth could be observed till 30th January 2020. From 30 Jan 2020 till 26th February 2020, 10% of total world population is under threat of COVID-19. This could be because of the initial reports (first report) of COVID-19 infection by less populated countries. The increase in COVID-19 cases showed linear growth from 26th Feb 2020 onwards. The data clearly shows that rate of virus spread increases rapidly with the increase in the number of diseased people. Further land area (KM^2) under threat of COVID-19 is analysed which is actually the land area of countries affected by SARS-CoV -2. This analysis will give us rough idea about the rate of spread of the virus.

3.2 Minimum Temperature and COVID-19 spread: The main aim of the present study is to explore the effect of temperature in the outbreak and spread of COVID-19. For this, we recorded the minimum temperatures (dated 23rd March 2020) of all those countries that reported COVID-19 infections. To deduce whether survivability of SARS-CoV-2 is dependent on environmental temperature or not we extrapolated the reported number of COVID-19 cases (till March 31, 2020) with respect to the minimum temperature of respective countries. The minimum temperature of most of the countries lies between -7.5^{o}C and 37.5^{o}C. Our results suggest the spread of COVID-19 infection in all those countries for which the minimum temperature lies between -7.3^{o}C to 30.2^{o}C. Majority of countries with reported COVID-19 cases have minimum temperature lying in the temperature interval of 22.7^{o}C to 30.2^{o}C, followed by 15.2^{o}C to 22.7^{o}C and 0.2^{o}C to 7.7^{o}C. Our data also suggested the most favoured temperature for the spread of COVID-19 to be between 14.5^{o}C and 29.30^{o}C.

To get more generalised idea of minimum temperature effecting the spread of COVID-19 infection, we first fit the distribution for the minimum temperature of infected countries and recorded that it follows t-distribution on the basis of minimum AIC (Akaike Information Criterion) and BIC (Bayesian Information Criterion). On the basis of estimated value of t-distribution, we fit the density plot into the histogram of minimum temperature with percentile chart of spreading the COVID-19 (Figure 4).

Figure-4

3.3 Maximum Temperature and COVID 19 spread: The information about the spread of COVID-19 with respect to maximum temperature is equally important. Hence we extrapolated the reported cases of COVID-19 infection in reference to maximum temperature (dated 23rd March 2020) of all the affected countries. Number of countries with the reported cases of COVID-19 in various intervals of maximum temperature is

represented in the form of percentile (Figure 5). The maximum temperature range for most of the countries lie between 1°C to 34°C, where as maximum temperature of almost 100 COVID-19 affected countries out of 202 countries lie between 17°C to 29°C. This suggests, temperature to be one of the important factors for the spread of SARS-CoV-2 and that the ambient temperature is well suited for the outbreak of COVID-19.

We also explored the maximum temperature recorded for all the countries on the basis of density plot (Figure 5). Best fitted model is Normal Distribution for the maximum temperature corresponding to minimum value of AIC and BIC, it shows that the nature of temperature is symmetrical for different countries.

Figure-5

3.4 Density of Infection: We explored the density of COVID-19 cases and deaths with respect to area of the country and its respective human population. Numbers of cases for 10,000 KM^2 land areas with recorded COVID-19 infection in descending order is shown in Figure-6. Maximum numbers of confirmed cases were recorded for USA followed by Italy. Similarly, countries like Monaco and Vatican with reference to land area, as they both are small countries. Maximum death density was recorded in Italy followed by Spain and the China. We also observed that till 31 March, 2020, Italy had reported more than 11 thousand deaths, which is 2.5 times more than the number of deaths reported in China (the epicenter of SARS-CoV-2 outbreak). The reason for huge gap between the numbers of deaths in Italy compared to China could be because of several factors such as internal health services. According to the percentile data the percent population of an affected country is more in San Marino followed by Vatican and Andorra.

Figure-6

To get more generalised idea about the COVID-19 out-break with respect to temperature, confirmed cases and deaths were analysed in the top 20 countries. Maximum and minimum temperature of these countries was recorded on 23 March, 2020. Maximum numbers of COVID-19 cases were reported in those countries where minimum and maximum temperature fell in the range of -30°C to 10°C and 15°C to 30°C, respectively. Maximum number of deaths was witnessed in those countries were minimum and maximum temperature fell in the range of -30°C to 10°C and 15°C to 30°C, respectively. Earlier it was observed that most of the affected countries have maximum temperature lying in the interval [17°C , 30°C] and the countries affected most have maximum temperature lying in [15°C , 30°C].

We further analysed the data of most infected cities with respect to their maximum and minimum temperatures. Majority of confirmed COVID-19 cases and deaths related to it were in those cities were minimum and maximum temperature fell in the range of 0°C to10°C and 0°C to 15°C, respectively. We also obtained the best fitted distribution on the basis of minimum and maximum temperature for the most infected cities and get log normal distribution is well fitted for this data set (Figure 7).

Figure 7

4. Distribution fitting on different class interval of maximum and minimum temperature: We analysed the best fitted distribution based on different temperature class intervals as discussed in section 3. For this, we considered the distributions such as Cauchy distribution, log-normal distribution, normal distribution, t- distribution, Weibull

distribution, exponential distribution, logistic distribution etc. This might be useful for describing the structure of the temperature interval because some of the distribution is explained in a bounded domain were as some in whole real line. In order to select the best particular distribution, information criterion such as AIC and BIC are considered to achieve the minimum value. We used R-software to find out the best distribution and based on estimated value of the best distribution we determined the density plot (Figure-8 a and b).

We observed that most of the temperature interval follows lognormal distribution as it attains minimum value of AIC and BIC. As it is known that lognormal distribution is positively skewed, the number of cases is going as high up to a certain interval and then it is decreased frequently. From this, we record the estimated value for the respective fitted distribution to get a generalized idea. For the given estimated values, we display the density plots for respective interval in Figure 8 (a) and Figure 8 (b).

Figure 8 (a)

Figure 8 (b).

5. Conclusion: The present paper analysed the effect of temperature in the spread of COVID-19 in different countries. The best fitted model based on minimum temperature follows a t-distribution with estimated parameters of 15.929 and 11.636, whereas maximum temperature follows a Normal Distribution with parameters of 21.95 and 9.181. After fitting the best distribution on temperature, the distribution of number of COVID-19 infected countries till 31 March 2020 was obtained for all class intervals recorded in the modeling of temperature. We recorded that the lognormal distribution fits well for COVID-19 total cases other than maximum infected countries belongs to the range of minimum temperature follows Weibull distribution. In the analysis of minimum temperature interval, the empirical average increases with the increase in temperature intervals. However, maximum variation is recorded in the interval [7.7,15.2]. It shows that spreading is more in the lower interval temperature however the variation is more in higher interval. Similarly, the empirical average is recorded in a decreasing manner. So, we conclude that the temperature plays a very important role in COVID -19, and infected countries with higher temperatures get a minimum of COVID 19 infections.

References:

1. Woo, P.C.; Huang, Y.; Lau, S.K.; Yuen, K.Y. Coronavirus genomics and bioinformatics analysis. *Viruses* **2010**, *2*, 1804–1820.
2. Zumla A, Chan JFW, Azhar EI, Hui DSC, Yuen KY. 2016. Coronaviruses-drug discovery and therapeutic options. Nat Rev Drug Discov 15:327–347.
3. Cui J, Li F, Shi ZL. 2019. Origin and evolution of pathogenic coronaviruses. Nat Rev Microbiol 17:181–192.
4. Zhong NS, Zheng BJ, Li YM, Poon LLM, Xie ZH, Chan KH, Li PH, Tan SY, Chang Q, Xie JP. 2003. Epidemiology and cause of severe acute respiratory syndrome (SARS) in Guangdong, People's Republic of China, in February, 2003. Lancet 362:1353–1358.

5. Zaki AM, Van Boheemen S, Bestebroer TM, Osterhaus ADME, Fouchier RAM. 2012. Isolation of a novel coronavirus from a man with pneumonia in Saudi Arabia. N Engl J Med 367:1814–1820.
6. Khan S, Nabi G, Han G, Siddique R, Lian S, Shi H, Bashir N, Ali A, Shereen MA. 2020. Novel coronavirus: how the things are in Wuhan. Clin Microbiol Infect.
7. Satu Hepojoki, Erika Lindh, Olli Vapalahti & Anita Huovilainen (2017) Prevalence and genetic diversity of coronaviruses in wild birds, Finland, Infection Ecology & Epidemiology, 7:1, DOI: 10.1080/20008686.2017.1408360
8. Zhu, N., Zhang, D., Wang, W., Li, X., Yang, B., Song, J., Zhao, X., Huang, B., Shi, W., Lu, R., et al.; China Novel Coronavirus Investigating and Research Team (2020). A Novel Coronavirus from Patients with Pneumonia in China, 2019. N. Engl. J. Med. https://doi.org/10.1056/NEJMoa2001017.
9. Pica N, Bouvier NM. Environmental factors affecting the transmission of respiratory viruses. Curr Opin Virol. 2012;2(1):90–95. doi:10.1016/j.coviro.2011.12.003
10. Linsey C. Marr , Julian W. Tang , Jennifer Van Mullekom and Seema S. Lakdawala, Mechanistic insights into the effect of humidity on airborne influenza virus survival, transmission and incidence, The Royal Society publishing, 16 January 2019
11. Anice C. Lowen, John Steel, Roles of Humidity and Temperature in Shaping Influenza Seasonality
12. Hampton T. Bats may be SARS reservoir. JAMA. 2005;294(18):2291. Banerjee A, Kulcsar K, Misra V, Frieman M, Mossman K. Bats and coronaviruses. Viruses. 2019;11(1):E41. https://doi.org/10.3390/v11010041.
13. Li W, Shi Z, Yu M, Ren W, Smith C, Epstein JH, et al. Bats are natural reservoirs of SARS-like coronaviruses. Science. 2005;310(5748):676–9.
14. Singh, A. K., Singh, A., & Engelhardt, M. (1997). The lognormal distribution in environmental applications. In *Technology Support Center Issue Paper*.
15. Delignette-Muller, M. L., & Dutang, C. (2015). fitdistrplus: An R package for fitting distributions. *Journal of statistical software*, *64*(4), 1-34.
16. Stedinger, J. R. (1980). Fitting log normal distributions to hydrologic data. *Water Resources Research*, *16*(3), 481-490.
17. Sánchez, S., Ancheyta, J., & McCaffrey, W. C. (2007). Comparison of probability distribution functions for fitting distillation curves of petroleum. *Energy & Fuels*, *21*(5), 2955-2963.