Special Issue: Pandemic Preparedness: Science and Countermeasures

Proceedings of the Pakistan Academy of Sciences:
B. Life and Environmental Sciences 58(S): 37-43 (2021)

ISSN (Print): 2518-4261; ISSN (Online): 2518-427X
http://doi.org/10.53560/PPASB(58-sp1)653

Pakistan Academy of Sciences

Research Article

Impact of Climate on Confirmed Cases of COVID-19 in Lahore: A Predictive Model

Farkhanda Manzoor, and Saima Sharif*

Department of Zoology, Lahore College for Women University, Lahore, Pakistan

Abstract: World over, the weather conditions are usually categorized as predictors of respiratory viral infections. This study uses a stepwise linear regression model to explore the effects of climate factors or weather factors such as temperature, humidity, wind speed on the spread of COVID-19 in Lahore city (Pakistan). The study was conducted in Lahore College for Women University, Lahore, and the data regarding the cases of COVID-19 in the Lahore district was obtained from the Primary and Secondary Health Care Department Punjab, from 18th March 2020 to 25th August 2020 while the weather statistics were obtained from Environmental Protection Department, Lahore. A predictive model by regression was designed in which day-to-day humidity, wind speed, average temperature, and their impact on confirmed cases of COVID-19 in Lahore were analyzed. The independent variables in the model were average temperature (C), humidity (%), wind (km/h) and the dependent variable was the number of daily established cases of COVID in Lahore. The result of the analysis shows the effectiveness of the proposed model and the impact of climate parameters on the assessment model. The study illustrates that the above model can be used to predict the future spread of COVID-19 based on the above-mentioned climate factors. As such, it proves as a useful modality to predict new cases for the government and other health agencies.

Keywords: COVID-19, climate, climate factors, stepwise linear regression, Lahore.

1. INTRODUCTION

The COVID-19 is an ongoing epidemic, caused by a coronavirus (nominated as severe acute respiratory syndrome coronavirus-2-SARS-CoV2) has been identified in December 2019 in Wuhan City, China. By rapidly spreading in too many countries, this virus has created a universal threat and was acknowledged as a "public health crisis of international significance" by the World Health Organization (WHO) on 30th January 2020. This virus spread from one person to another via indirect and direct physical contact with contaminated objects or surfaces besides open places with aerosols and aerial droplets generated through speaking, coughing, or sneezing [1, 2].

Numerous studies have been conducted in different countries to monitor this pandemic to strengthen their understanding of the factors that affect viral transmission (including climate and weather). Respiratory viral infectious diseases like corona-virus, Severe Acute Respiratory Syndrome (SARS) and seasonal influenza tends to increase in the colder environment, thereby weather conditions are considered as one of the factors for acquiring respiratory infections [3-6].

Although seasonal influenza and SARS have some similarities to the new coronavirus, it has been hypothesized that its existence may depend on the weather and assumed that low temperature tends to promote the escalation of the virus [6,7]. In comparison, humid and hot weather has been speculated to minimize viral activity [8]. This remains uncertain as there is no consensus to date, on the effect of humidity and temperature to regulate COVID-19 dynamics or tropical-humid environments overwhelm its transmission. Globally, several studies investigated the outbreak of COVID-19 in terms of different weather conditions, especially humidity and temperature [9,11]. These studies have described negative or positive associations among the transmission of COVID-19

Received: March 2021; Accepted: July 2021
*Corresponding Author: Saima Sharif <ssharif1978@yahoo.com>

and changes in humidity and temperature. This study aims to discover the consequence of climatic conditions on the prevalence of COVID-19 in Lahore (Pakistan) and to propose a predictive model to improve the prediction.

2. MATERIALS AND METHODS

2.1 Data Set

The data regarding established cases of COVID-19 in the Lahore district were obtained from the Primary and Secondary Health Care Department, Punjab, for the period from 18th March 2020 to 25th August 2020. The data of climatic or weather parameters were obtained from the Environmental Protection Department, Lahore. As the levels of humidity, temperature, and wind speed change during the day, the average temperature, humidity, and wind speed were noted each day for the period from 18th March 2020 to 25th August 2020. The MATLAB software was used to form a predictive model to determine the role of climate on the outburst or spread of COVID-19. A regression model was used for this purpose. Data was analyzed to finalize which predictors should be used in the final regression model. This process was referred to as the variable selection problem. Two objectives were involved in determining these subsets of independent variables. The first objective was to make the model thorough and as precise as possible i.e., to include every regressor that was even slightly related to the dependent variable. Simultaneously, the least number of independent variables were included in the model because irrelative predictors make the model imprecise, complex thereby making the processor of variable selection a conundrum: striking a balance between fit (as many predictors as possible) and simplicity (to include a minimum number of regressors).

2.2 Screening of Candidate Variables for the Model

When a large group of variables occurs, the candidate variables for use in the Model may be screened from the group of variables by using the techniques namely, the Backward (Step-Down) Selection Procedure, or the Forward (Step-up) Selection Procedure [12, 13]. The Backward (Step-Down) Selection method is less popular because it begins with a model in which all candidate variables have been included. However, because it works its way down instead of up, it always retaining a greater value of R-Squared. The problem is that the models selected by this mode may contain variables that are not required. The user sets the effect level upon which variables can pass in the model. The backward selection model begins with variables used for all candidates in the model. At every phase, the least important variable is removed until no non-significant variables remain. The significance level at which variables can be removed from the model is set by the user.

The Forward (Step-up) Selection Technique is frequently used to provide a preliminary screening of the candidate variables when a large group of variables occurs [12]. A reasonable approach would be to use the forward selection procedure to obtain the best ten to fifteen variables and then apply the all-possible algorithm to the variables in this subset. This procedure is also a good choice when multicollinearity is a problem. The forward selection method is simple to define. It begins with no candidate variables in the model. A variable that has the highest R-Squared is selected. At each step, only that candidate variable is selected in which R-Squared raises the most. When none of the remaining variables are major then the process of adding variables is stopped. This exercise of adding variables in the model is practiced by taking into account the fact that once a variable enters the model, it cannot be deleted.

P-value is the measure to judge the result of the model. P-value is essentially the probability of having results that are almost at the extreme as the actual results, assuming the null hypothesis is correct.

F-test is done to check if the group of the independent variable are collectively significant or not. F-value shows whether a regression model gives better-fitted data than a model having no independent variable. Similarly, the p-value and F-value are the measures used to check the effectiveness of the constructed model. The smaller values mean that extreme outcomes are highly unlikely to occur.

2.3 Stepwise Regression

Stepwise regression is a logical mode for removing and adding parameters from a generalized linear or linear model based upon the statistical importance in elucidation of the response variable. The technique initiates with a primary model and then relies on the descriptive power of incrementally smaller and larger models. The stepwise regression uses backward and forwards regression to determine the last model. At every phase, the function pursuits for terms to remove from the model or add to the model based on the p-value for the F-test of the change in the sum of squared error that results from adding or removing the term. Stepwise regression takes the following steps: The initial model was fitted.

1. Examine a set of available terms not in the model. If any of the terms have p-values less than an entrance tolerance (i.e. if it is unlikely a term would have a zero coefficient is added to the model), add the term with the smallest p-value and repeat this step; otherwise, go to step 3.
2. If any of the available terms in the model have p-values greater than an exit tolerance (that is, the hypothesis of a zero coefficient cannot be rejected), remove the term with the largest p-value and return to step 2; otherwise, end the process.

3. RESULTS

In the present study, weather factors (average wind speed, average temperature, and average humidity of each day) and COVID -19 data of each day for the period from March 2020 till August 2020 were used. Figures 1, 2 & 3 showed the number of COVID-19 cases with average temperature (^{o}C), average humidity (%), and average wind speed (km/h) respectively in the Lahore district. In the regression model, the independent variable was Average Temperature (^{o}C), Humidity (%), Wind (km/h), and the dependent variable was daily confirmed cases of COVID in Lahore.

Linear regression model: COVID-CASES ~ 1 + Humidity + Wind*Temperature

The constructed linear regression model was presented in Table 1. ANOVA test was performed on the constructed model which gave the results as presented in Table. 2. The values of R^2 were at the appropriate level, and the regression model was significant since the final model p-value is 0.02361 which is less than 0.05. The regression model plot as shown in Figure 4 illustrates that the model was significant because a horizontal line does not fit between the confidence bounds, which is consistent with the p-value obtained. The residuals plot of the linear regression model is shown in Figure 5.

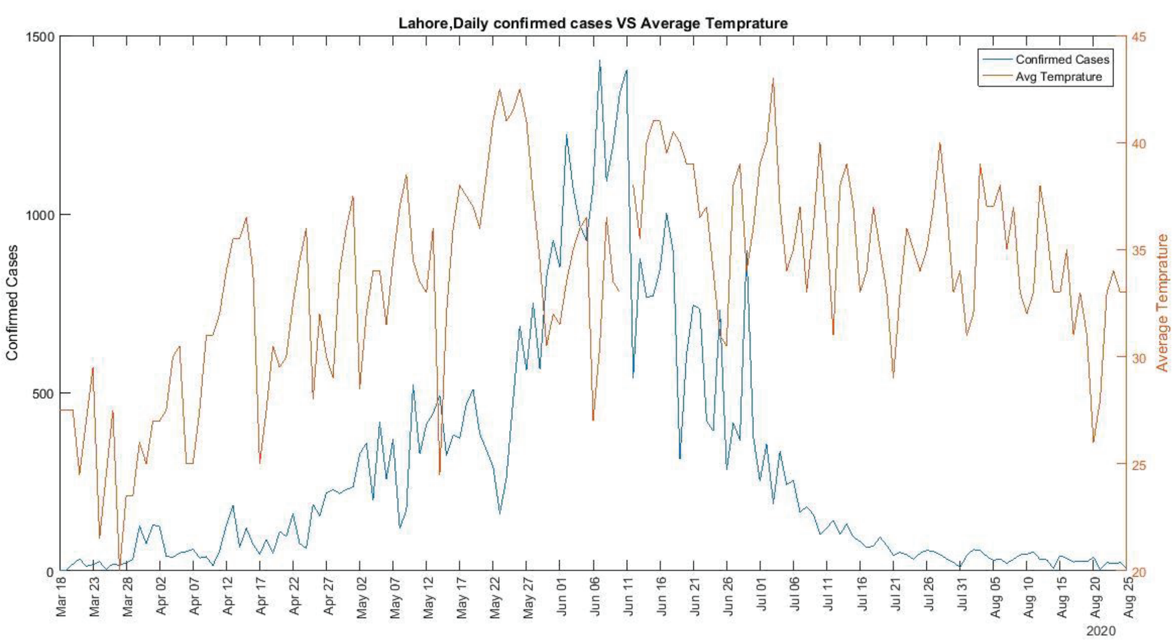

Fig. 1. The average temperature during the study period in Lahore

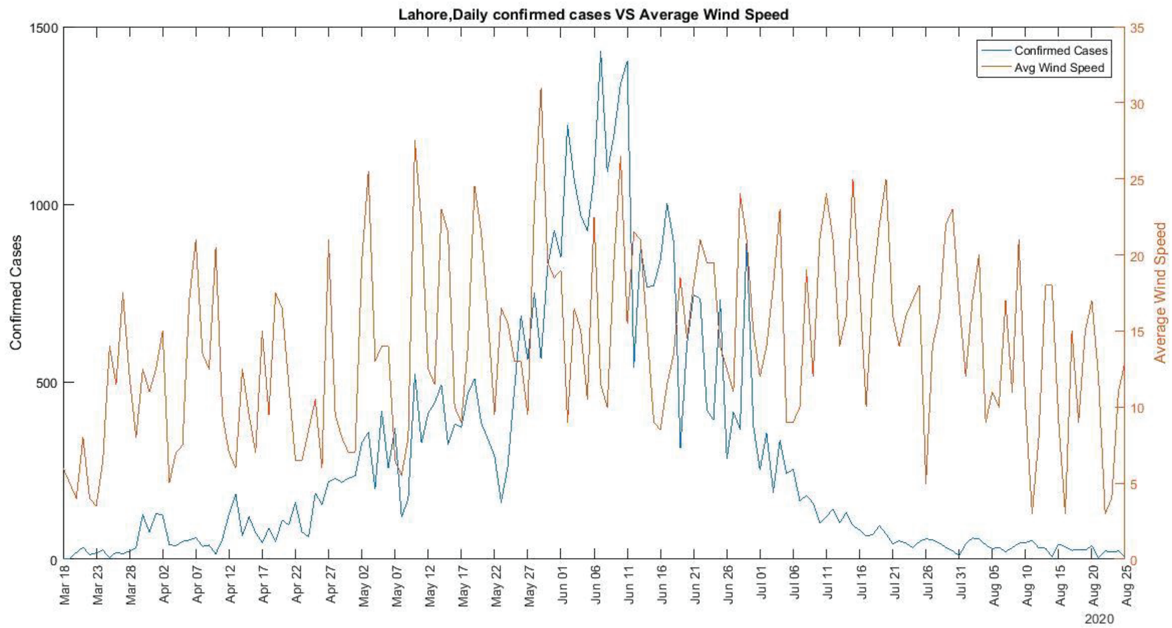

Fig. 2. The average wind speed during the study period in Lahore.

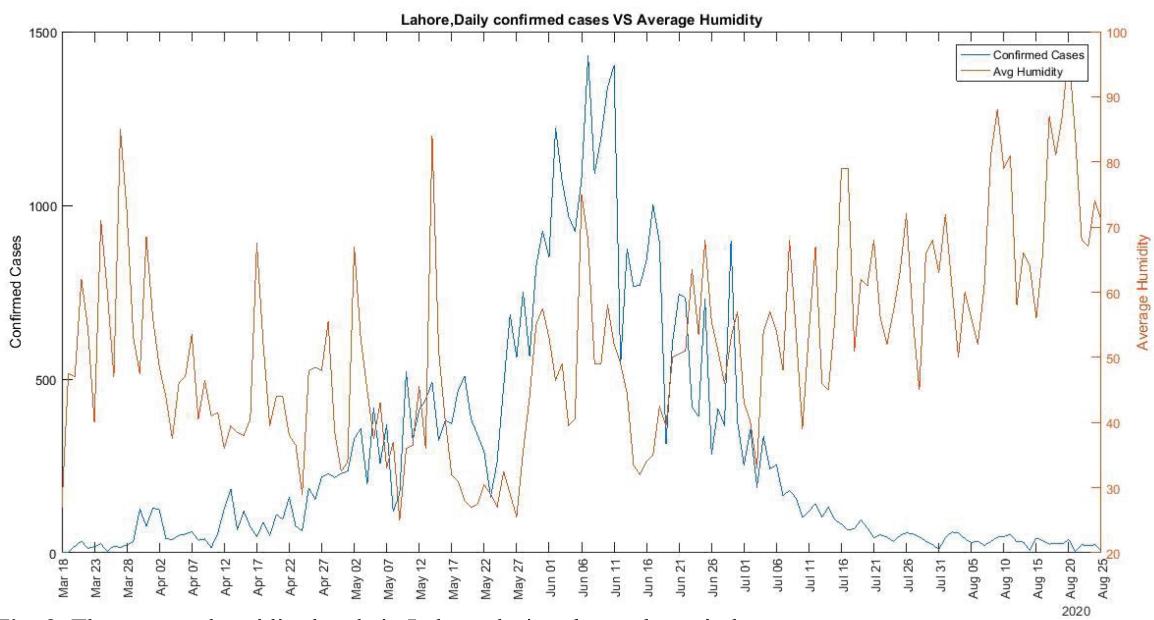

Fig. 3. The average humidity levels in Lahore during the study period.

Table 1. Linear Regression Model Estimated Coefficients:

	Estimate	**SE**	**T Stat**	**P -Value**
(Intercept)	-1082.8	495.57	-2.1849	0.031
Wind	91.921	34.082	2.6971	0.007
Humidity	-4.5781	1.7713	-2.5847	0.011
Temperature	41.669	14.477	2.8784	0.005
Wind: Temperature	-2.3094	1.0103	-2.2858	0.024

R-squared: 0.187, Adjusted R-Squared 0.166

F-statistic vs. constant model: 8.91, p-value = 1.67e-06

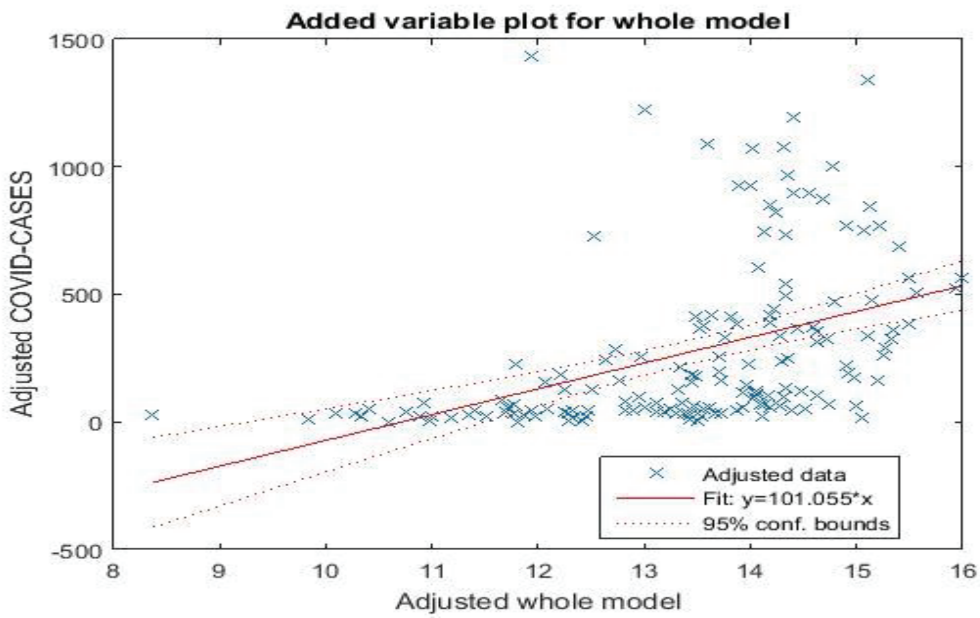

Fig. 4. Regression model plot between adjusted climate parameters and COVID cases

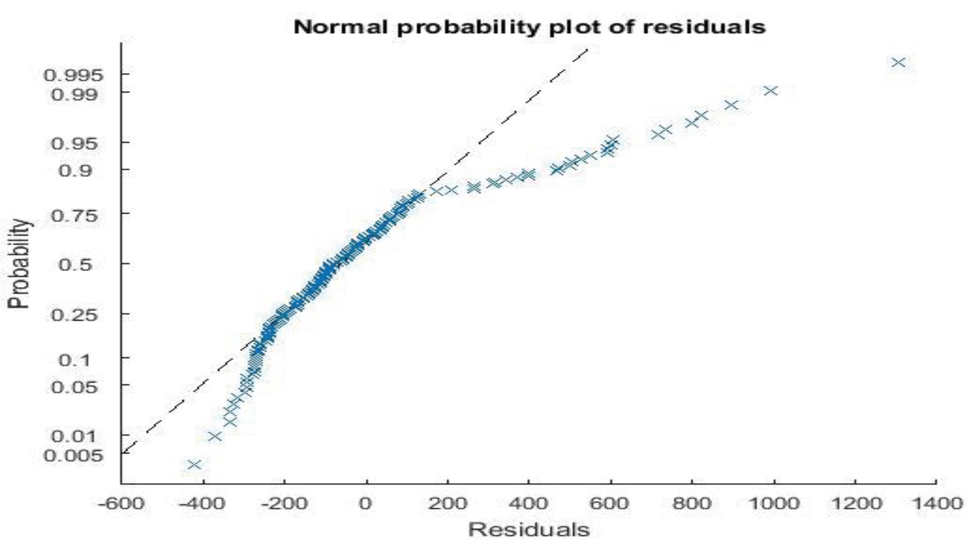

Fig. 5. Residuals of Linear regression.

Table 2. ANOVA Analysis

	Sum Sq	D F	Mean Sq	F	P- Value
Wind	1.1463e+06	1	1.1463e+06	13.004	0.001
Humidity	5.8885e+05	1	5.8885e+05	6.6804	0.011
Temperature	3.3949e+05	1	3.3949e+05	3.8515	0.051
Wind : Temperature	4.6057e+05	1	4.6057e+05	5.225	0.024
Error	1.3663e+07	155	88146		

4. DISCUSSION

In China's 130 cities (except Wuhan), it was reported that in the initial stage of the epidemic, the spread of COVID-19 was due to humidity and temperature [9]. In the same way, the facts of other countries, excluding China revealed that everyday COVID-19 new cases were inversely related to temperature and relative humidity [9], while the mean day to day temperature had a major effect on the daily reported COVID-19 cases. For the spreading of SARS-CoV-2, the optimum temperature was 8.07 °C and humidity varied from 60% to ~90%. This is in accordance with Wang et al [8] that who emphasized the role of temperature and reported that 8.72°C was the optimum temperature for viral spread. The decrease of temperature and humidity tends to increase the virus survival and spread while the study by Rouen *et al*.,(2020) recorded a negative relationship between the daily temperature and prevalence of COVID-19 [14], a study by Rasheed *et al*., (2020) reveal that a climatic aspect, especially temperature, is a noteworthy environmental issue for the dispersion of COVID-19 infection [15]. Reportedly, COVID-19 infection was primarily more associated with temperature between 5°C to 11°C and absolute humidity levels less than three [10].

A positive association between COVID-19 prevalence and increases in temperature and humidity has also been documented [3, 16, 17]. In Singapore, positive associations with the amount of regular as well as aggregate COVID-19 cases were related to absolute humidity and temperature [3]. A direct association was observed during the preliminary stage of the COVID-19 outburst and ambient temperature. The mainstream of the reviewed studies indicates a stimulating trend. In several studies, the important negative association of COVID-19 cases with humidity and temperature suggests that SARS-CoV-2 has a periodic impact, with dry and cold circumstances, which can promote the spread of the new coronavirus, while at higher temperatures and higher comparative humidity levels, the virus loses its capability and thus weakens its spread [17, 18]. The results were consistent with reports of SARS-CoV, influenza, and other corona-viruses, i.e., Extreme Acute Respiratory Syndrome [19]. The occurrence of COVID-19 may, therefore, verify to be seasonal, which presents a warning in the winter season for increased transmission. Certain studies have also revealed that COVID-19's physical distribution has been confined to countries with colder climates so far [20]. In the low-temperature area of China, this virus first appeared in early winter, and subsequently, the main epidemic happened in further Asian states such as Middle East, Europe, Japan, and South Korea where during January and February 2020, the temperatures were far below 0°C [21, 22]. As the virus spread, a narrow climatic band was initially grouped into new epicenters. Cold weather tends to be further conducive to the transmission of COVID-19 that elucidate the consistently higher COVID-19 cases in these countries.

5. CONCLUSION

This study indicates the use of a stepwise linear regression model to check the effects of climate factors such as temperature, humidity, wind speed on the spread of COVID-19. The model is evaluated based on p-value, Root mean square value, and R squared value. The results show that the above model can be used to predict the future outbreak of COVID-19 based upon the above-mentioned climatic factors. It could prove useful to the primary and secondary health care departments, clinical/hospital management, as well as pandemic prevention-related agencies of the Government in preventing the virus. The study can be further extended by using the other regression models and deep learning methods.

6. ACKNOWLEDGEMENTS

The authors acknowledge the cooperation by the Primary and Secondary Health Care department as well as of Environmental Protection Department, Punjab.

7. CONFLICT OF INTEREST

The authors declare no conflict of interest.

8. REFERENCES

1. G. L. Rosa, L. Bonadonna, L. Lucentini, S. Kenmoe, and E. Suffredini. "Coronavirus in water environments: Occurrence, persistence and concentration methods-A scoping review." *Water research* :115899 (2020).

2. Q. Li, X Guan., P. Wu., X. Wang., L. Zhou., Y. Tong, Ruiqi Ren et al. "Early transmission dynamics in Wuhan, China, of novel coronavirus–infected pneumonia." *New England journal of medicine* (2020).
3. S. K. Pani., N. H. Lin, and S. RavindraBabu. "Association of COVID-19 pandemic with meteorological parameters over Singapore." *Science of the Total Environment.* 740: 140112 (2020).
4. P. K. Sahoo., S. Mangla., A. K. Pathak., G. N. Salãmao, and D. Sarkar. "Pre-to-post lockdown impact on air quality and the role of environmental factors in spreading the COVID-19 cases-a study from a worst-hit state of India." *International journal of biometeorology:* 1-18 (2020).
5. S. Hamidi., S. Sabouri, and R. Ewing. "Does density aggravate the COVID-19 pandemic? Early findings and lessons for planners." *Journal of the American Planning Association* 86, no. 4 :495-509 (2020).
6. P. Mecenas., R. T. d. R. M. Bastos., A. C. R. Vallinoto, and D. Normando. "Effects of temperature and humidity on the spread of COVID-19: A systematic review." PLoS one 15, no. 9 : e0238339 (2020).
7. Y. Ma., Y. Zhao., J. Liu., X. He., B. Wang., S. Fu., J. Yan., J. Niu., J. Zhou, and B. Luo. "Effects of temperature variation and humidity on the death of COVID-19 in Wuhan, China." *Science of the total environment* 724 : 138226 (2020).
8. J. Wang., K. Tang., K. Feng, and W. Lv. "High temperature and high humidity reduce the transmission of COVID-19." Available at SSRN *3551767* (2020).
9. Y. Wu., W. Jing., J. Liu., Q. Ma., J. Yuan., Y. Wang., M. Du, and M. Liu. "Effects of temperature and humidity on the daily new cases and new deaths of COVID-19 in 166 countries." *Science of the Total Environment* 729 :139051 (2020).
10. M. M. Sajadi., P. Habibzadeh., A. Vintzileos., S. Shokouhi., F. Miralles-Wilhelm, and A. Amoroso. "Temperature, humidity, and latitude analysis to estimate potential spread and seasonality of coronavirus disease 2019 (COVID-19)." *JAMA network open* 3, no. 6: e2011834-e2011834 (2020).
11. P. Mecenas, R. T. D. R. M. Bastos, A. C. R. Vallinoto, and D. Normando. Effects of temperature and humidity on the spread of COVID-19. A systematic review. *PLoS one* 15: e0238339 (2020).
12. K. Wang, and Z. Chen. Stepwise Regression and All Possible Subsets Regression in Education. *Electronic International Journal of Education, Arts and Science* 2: 60-81 (2016).
13. Y. Kim. Toward a successful CRM: variable selection, sampling, and ensemble. *Decision Support Systems,* 41 (2), 542-553, (2006).
14. A. Rouen., J. Adda, O. Roy, E. Rogers, and P. Lévy. COVID-19: relationship between atmospheric temperature and daily new cases growth rate. *Epidemiology & Infection* :148 (2020).
15. E. A. Rasheed, S. Kodera, J. Gomez-Tames, and A. Hirata. Influence of Absolute Humidity, Temperature and Population Density on COVID-19 Spread and Decay Durations: Multi-Prefecture Study in Japan. *Int. J. Environ. Res. Public Health,* 17, 5354. (2020).
16. R. Basray, A. Malik, W. Waqar, A. Chaudhry, M. W. Malik, M. A. Khan, J. A. Ansari, and A. Ikram. Impact of environmental factors on COVID-19 cases and mortalities in major cities of Pakistan. *Journal of Biosafety and Biosecurity* 3: 10-16 (2021).
17. K. H. Chan., J.S. M. Peiris., S. Y. Lam., L. L. M. Poon., K. Y. Yuen, and W. H. Seto. "The effects of temperature and relative humidity on the viability of the SARS coronavirus." *Advances in virology* 2011 (2011).
18. L. Hu, W. J. Deng, G. G. Ying, and H. Hong. Environmental perspective of COVID-19: atmospheric and wastewater environment in relation to pandemic. *Ecotoxicology and Environmental Safety* : 112297 (2021).
19. M. R. M. Daneshvar, M. Ebrahimi, A. Sadeghi, and A. Mahmoudzadeh. Climate effects on the COVID-19 outbreak: a comparative analysis between the UAE and Switzerland. *Modeling Earth Systems and Environment:* 1-14 (2021).
20. M. M. Iqbal, I. Abid, S. Hussain, N. Shahzad, M. S. Waqas, and M. J. Iqbal. The effects of regional climatic condition on the spread of COVID-19 at global scale. *Science of the Total Environment* 739: 140101 (2020).
21. T. Jamil., I. Alam., T. Gojobori, and C. M. Duarte. "No evidence for temperature-dependence of the COVID-19 epidemic." *Frontiers in public health* 8 : 436 (2020).
22. G. Kroumpouzos., M. Gupta., M. Jafferany., T. Lotti., R. Sadoughifar., Z. Sitkowska, and M. Goldust. "COVID-19: A relationship to climate and environmental conditions?" *Dermatologic therapy* (2020).