International Journal of Computing and Digital Systems
ISSN (2210-142X)
Int. J. Com. Dig. Sys. 11 , No.1 (Feb-2022)

https://dx.doi.org/10.12785/ijcds/110174

Impact of Weather on COVID-19 in Metropolitan Cities of Pakistan: A Data-Driven Approach

Mehak Narmeen[1], Mian Usman Sattar[2,3], Hamza Wazir Khan[1], Mehak Fatima[1], Mueen-ud-Din Azad[4] and Fatima Ghani[1]

[1]*Department of Information Systems, University of Management and Technology, Lahore, Pakistan.*
[2]*Department of Management, School of Business, Beaconhouse National University, Lahore, Pakistan*
[3]*Department of Information Technology, School of Science and Engineering, Malaysia University of Science and Technology, Petaling Jaya, Selangor, Malaysia*
[4]*Department of Economics, University of Management and Technology, Lahore, Pakistan.*

E-mail address: f2017313005@umt.edu.pk, usman.sattar@bnu.edu.pk , hamza.khan@umt.edu.pk, f2017313027@umt.edu.pk, mueen.azad@umt.edu.pk, f201831303020@umt.edu.pk

Received 7 Aug.2020 , Revised 3 Jul.2021 , Accepted 8 Jul. 2021 , Published 15 Jan.2022

Abstract: The world has been drastically affected by the 2019 novel corona-virus outbreak, otherwise known as COVID-19. Pakistan is no exception to this, and the virus has rapidly spread within the country in the first quarter of 2020. As previously some of the contagious diseases such as influenza have been affected by various meteorological factors therefore, several recent studies have been done in various parts of the world to find the climatic impact on COVID-19 as well. This study also, analyzes the effect of meteorological factors on the proliferation of COVID-19 in Pakistani urban centers, namely, Islamabad, Karachi, Lahore, and Peshawar. The variables observed for this purpose are the number of confirmed cases, fatalities, recoveries, and transmission type. Finally, we also forecast the trend of COVID-19 in the coming days, using a time series model named as Facebook's Prophet Library.

Keywords: COVID-19; corona-virus diseases; Facebook Prophet Library; Weather Forecast; Pakistan

1. Introduction

Throughout history, plagues, epidemics, and pandemics have withered humanity and shaped the upcoming future. A pandemic is a disease that spread throughout the continents and over multiple countries. These are not new to the world and go back to the beginning of time. Human history has recorded several pandemics of diseases such as smallpox and tuberculosis [1] The most lethal of which recorded history to date was the Black Death, more commonly referred to as the great plague, which killed approximately 75–200 million people in the world in the 14th century [2]. While some of these recorded pandemics were not transmitted through direct human contact, the deadliest of these diseases were contagious, resulting in mass deaths in the infected areas.

In the 21st century, humans were affected by two highly transmissible and pathogenic viruses named Middle East respiratory syndrome corona-virus (MERS-CoV) and severe acute respiratory syndrome corona virus (SARS-CoV). Both these viruses prospectively originated in bats. Additionally, genetically diverse corona-viruses that fall under the umbrella of SARS-CoV and MERS-CoV were globally found in bats species.

Recently the outbreak of severe acute respiratory syndrome corona-virus 2 (SARS-CoV-2; previously provisionally named 2019 novel corona-virus or 2019-nCoV) disease (COVID-19) has created mayhem in the world. Corona-viruses come from a large family of viruses, and it is observed that different classes of the virus cause illness in humans. At the same time, some circulate only in animals, including cats and bats. On December 31, 2019, the first case of the novel Corona Virus was reported in Wuhan, China, by the World Health Organization (W.H.O) from its country office in China [3]. WHO declared this outbreak a Public Health Emergency of International Concern on January 30, 2020, and later a pandemic on March 11, 2020 [3], [4]. The novel COVID-19 is a respiratory illness spread through human interaction and exposure to infected surfaces. It usually causes flu-like symptoms at first and can also result in severe respiratory discomfort leading to death. As of June 3, 2020, it has killed around 380,662 people globally [5]. With no proper and effective vaccine

E-mail address: f2017313005@umt.edu.pk, usman.sattar@bnu.edu.pk, hamza.khan@umt.edu.pk, f2017313027@umt.edu.pk, mueen.azad@umt.edu.pk, f201831303020@umt.edu.pk

or treatment in place, there are several suggested strategies and standard operating procedures (S.O.P.) set in place by global health organizations, notably WHO. These strategies include social distancing, focus on hygiene, quarantine, and strengthening immunity.

In the absence of vaccines and medicines, when quarantine sounds to be the only way of reducing the spread of this disease, it is essential to gather factors impacting its outbreak. Scientists are trying to point out these factors for understanding the nature of this infection. Meteorological factors were said to be influencing the viral spread of COVID-19. Initially, when the outbreak took place, it was noted that weather and environmental factors affect its spread [6]. Worldwide data was also used to investigate the impact of the increase in temperature in restricting the virus's spread. However, nothing has been proved certain yet [7]. With every passing day, some new aspect of this viral disease comes out. To reach out to some certain conclusive aspects, studies have been done by scientists of many fields concerning the different aspects of the disease [8], [9].

Likewise, computer scientists are trying to find out various factors that impact this pandemic [8], [10]. A recent study has been done to explore and analyze the main issues in forecasting epidemics like dengue, using data science processes [11]. The main objective of this study, in particular, is to scientifically answer the question of whether the change in weather affects COVID-19 or not, using a data-driven approach. It is not simple to answer the question because of the involvement of various differentiating environmental factors, along with the weather.

This paper is divided into sections; section II gives a brief review of the research done related to this area. As this is a recent outbreak, not many studies have been done until now. Section III includes details of data sources and collection and a brief description of the methodology. The methodology consists of LASSO (Least Absolute Shrinkage and Selection Operator) feature selection to find the importance of the weather components and multiple regression analysis for determining the impact of variables on the model. Furthermore, Facebook's Prophet Library to forecast the overall spread. Section IV focused on discussing the results. Lastly, the conclusion is given in Section V of the paper.

2. Related Work

Many considerations can be taken in view of this virus that can contribute to the increased mortality rates. It is important to keep in mind the relative newness of this disease and the minimal and somewhat premature stages of research that has gone into determining the factors that can be involved in the spread of the virus. However, the global indicators of COVID-19 have allowed some domineering factors to surface in many studies. For example, in a study done by Lakshmi and Suresh in 2020 through a TISM approach, it was found that changes in interlinked factors such as the Humidity of the environment, hygiene practices, potency of the virus, population density, airflow and ventilation of living spaces, etc. cause a direct and indirect effect in the spread of the viral [12]. Overall, the study concluded that elements like social distancing, age, air temperature, airflow, and ventilation are driving factors that cause rapid transmission and are interdependent in the spread of the disease [13].

Furthermore, a study published in China C.D.C. Weekly discusses the impact of airflow and ventilation on the spread of the virus. The study focuses on a descriptive analysis of federal cases in China during the peak time – December till February 2020. The study explored how close areas with low airflow and ventilation, among other things, can have a direct impact on the risk of being infected by COVID-19. Much like influenza's' such as flu and rhinovirus, COVID-19 was believed to be spread through respiratory droplets released in the air through coughing/sneezing [14]. These droplets would infect others through aerosol transmission and prolonged contact with elevated aerosol concentrations in closed spaces. People within six feet of ejected respiratory droplets are likely to be exposed to them and thus can contract the disease [15] [16]. The study also indicated that close contact between the host and the individual was necessary for the spread of the disease along with contact with surfaces touched or used by the host body.

Additionally, the evidence also suggested that the health and age factors were also directly linked to the possibility of being infected by the disease. People with compromised immune systems and aged individuals were at high risk. Shared ventilation ducts were also seen as a possible risk factor.

A study was done by Onder, Rezza, and Brusaferro in 2020 examined evidence from various countries globally and concluded that COVID-19 fatally impacts older patients. Data taken from quantitative measures indicated that case-fatalities of patients from the age groups 70-80 years were relatively high in all countries. They believe that these fatalities are not merely a result of physical factors; there is also an underlying social factor contributing to the high mortality rate in the older generation.

Firstly, the older generation has weak immune systems and thus is more prone to attacks by any virus, not just COVID-19. Any prevailing health conditions, such as diabetes or respiratory illnesses, make them vulnerable by lowering the overall immunity of the body. The added stress of suffering from the COVID-19 virus also further compromises the immune system and the ability of the body to fight back in response. Secondly, many elderly members at the ages of 70-80 suffer from mobility challenges and thus are unable to isolate or quarantine themselves, which is why they are at more chance of contracting the disease from the host body [17]. Furthermore, elderly members of societies are often living in nursing or retirement homes and

are cooped up in one shared space [18].

According to Lewnard and Lo, social distancing is one of the most successful non-pharmaceutical prevention strategies introduced to ensure that the spread of the disease is minimized. With disease control being declared a top priority by WHO, public health authorities globally have emphasized on S.O.P.s and social distancing. Furthermore, legalizations and laws have been introduced to ensure that individuals thoroughly follow social distancing along with other protective measures. Countries all over the globe have enforced community-level lock downs and social distancing to minimize the spread of COVID-19. Other controlling efforts, such as contact tracing, are also being administered. The researchers state that social distancing measures if implemented timely and efficiently, slow down the spread of the virus [19]. However, the research also points out that the effectiveness and social impact of measures such as isolation and quarantine depend heavily on the approaches taken by public health authorities in the country along with political leaders and institutions such as media [20]. A study done on the spread of influenza in China through different modes of transport suggests that air and train travel must be limited during the early outbreak as they are key risks in the transmission, while highways should also be considered important equally [21].

Another study was done by Jia, Yuan, Xu, Jia, and Christakis in 2020 looked into how rapid, significant, and diffused human migration has amplified localized outbreaks of the virus into a pandemic. They note that this was also the case with COVID-19. They study the spread of the virus by tracking the disease's movement with real-time data focused on aggregate population flows [22]. Study done by used nonlinear delay dynamical modelling to study this pandemic [23].

Despite having been studied extensively from qualitative and quantitative aspects, the COVID-19 virus remains a mystery when it comes to spreading and prevention analysis. The researches mentioned above have focused on multiple aspects that play a role in the spread of the virus or on solely one factor through qualitative and quantitative interpretations. None of the researches mentioned above has utilized advanced machine learning strategies to isolate how the weather impacts the spread of COVID-19 in countries that experience a wide range of weather conditions in different regions. Pakistan is one such country [24].

In this paper, we will be isolating the impact of weather on COVID-19 through different methodologies with a sole focus on the spread of the virus in Pakistan.

3. METHODS AND MATERIALS

A. Data and Sources

The data used for modeling and prediction analysis includes numbers regarding confirmed cases, deaths, and recoveries in Pakistan. Various sources maintain the data, but the one used in this paper is the default one managed by the Johns Hopkins University Center for Systems Science and Engineering (JHU CSSE). This data is made available to the public through a COVID-19 Python library. City-wise COVID-19 data-set for Pakistan is acquired by Kaggle. The data consisted of numbers indicating confirmed cases, recoveries, deaths, travel history of those cases, and location [4]. Moreover, the historical weather data was collected using Python's library wwo-hist. The weather data consisted of multiple weather components, including Maximum Temperature, Minimum Temperature, U.V. Index, Dew-point, Heat Index, Wind Gust, Humidity, Precipitation, Pressure, etc.

B. Data Preprocessing

The data collected from the COVID-19 library was already preprocessed. However, data obtained from Kaggle needed preprocessing. The data was checked and corrected for any inconsistencies in the names of cities, provinces, and travel history, and uniqueness was maintained. Coordinates for the cities were taken from Google. The weather data was then conjoined to create a single data frame with the city-wise data based on the date and location using coordinates.

C. Feature Selection

Since the data shows multicollinearity between the input variables, therefore, we have used LASSO Regression for feature selection (reduce the number of predictors while developing a predictive model). Lasso uses L1 norm as a regularize. It derives parameters to zero. Higher the value of alpha, the fewer features have non-zero values.

D. Forecasting

In this study, we have used Facebook's Prophet Library to forecast the spread (number of cases) of COVID-19 in Pakistan (includes all cities). The range depends on the number of conditions and can only be estimated if the conditions remain the same.

Time series data has been forecasted by using Facebook's Prophet. It is based on an additive model where non-linear trends are fit with seasonality (yearly, weekly, and daily), plus holiday effects.

$$y(t) = g(t) + s(t) + h(t) + \epsilon t \quad (1)$$

g(t): linear or logistic growth curve for modeling non-periodic changes in time series

s(t): seasonal changes that can include daily/weekly/yearly seasonality

h(t): effect of holidays (can be irregular schedules)

ϵt : error term that balances the model for any unusual changes not accommodated by the model

Facebook's Prophet Library is an excellent technique for time series analysis and has been used to predict the number of upcoming cases in this study. The Prophet's algorithm

takes only two columns as an input (y – Target (Cases) and ds Date time).

4. Results

The data set contains several confirmed cases, fatalities, and recoveries reported in different cities of Pakistan. Therefore, we have divided the results among the four major cities of Pakistan. These cities include Islamabad (Islamabad Capital Territory), Lahore (Capital of Punjab province), Karachi (Capital of Sindh province) and Peshawar (Khyber Pakhtunkhwa). Since these are major metropolitan cities in Pakistan, it highly justifies the higher number of cases in these cities. Also, the healthcare facilities are much better and higher in number in these cities than the other cities of Pakistan.

	Province	Cases	Deaths	Recovered
0	Azad Jummu Kashmir	285	6	110
1	Baluchistan	5125	49	1981
2	Federal Administration Tribal Area	53	1	15
3	Gilgit-Baltistan	787	24	363
4	Islamabad Capital Territory	3523	36	460
5	Khyber Pakhtunkhwa	10259	473	1939
6	Punjab	31096	593	4081
7	Sindh	32858	546	15805

Figure 1. Province wise analysis of the number of COVID-19 Cases, Deaths and Recoveries

Figure 1. shows the overall number of cases in different provinces. The rise of cases can be due to multiple reasons, but as this paper focuses on determining the impact of temperature, we have only considered meteorological factors for the above-mentioned places. Most number of cases are in Sindh followed by Punjab while the greatest number of deaths are recorded in Punjab till date. Also, the recovery rate is much higher in Sindh than any other province.

TABLE I. Population in Cities of Pakistan

City	Province	Population
Islamabad	Islamabad Capital Territory	2,001,579
Karachi	Sindh	11,589,068
Lahore	Punjab	11,126,285
Peshawar	Khyber Pakhtunkhwa	4,269,079

Table 1. indicates the population in capital cities of these provinces [25]. Karachi is the most populated city of Pakistan followed by Lahore.

A. *Measuring the impact of Temperature*

Pakistan has extremely varying temperatures as it lies above the tropic of cancer [26]. The northern part of the country remains cool throughout the year, while the south has moderate to extreme temperatures. Since the start of the pandemic, the overall temperatures in the majority of the areas under focus were cold.

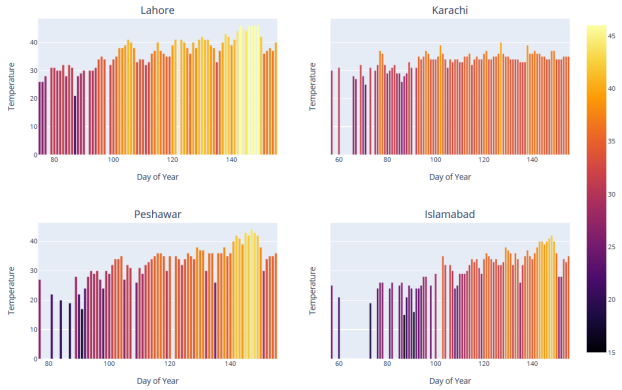

Figure 2. Shows the temperatures in the cities starting from February 26, 2020 till June 3, 2020

Figure 2. shows that as time progressed; the temperatures began rising in different cities of Pakistan. The temperatures are shown from February 26, 2020, as the country's first case was reported on that date, while June 3, 2020, is taken as a cutoff date (data concluded).

The temperature scale shows Lahore, Peshawar, and Islamabad had cold temperatures that progressed to high temperatures while Karachi has moderate temperatures during this period.

The primary reason for selecting multiple regions for this study is to include all types of temperature variations. As a result, the data shows a variation of various temperature levels in different areas of the country, which increases the reliability and strength of our sample data and is enough to conclude the results below.

Since the start of this pandemic, Lahore has seen a steady increase in the number of Corona cases. As shown in the Fig 3. below, a sudden spike from week 16, on-wards can be seen in the cases.However, the recoveries also started to jump up from week 20 while the number of fatalities remains steady till week 22.

Karachi has also seen a steady increase in the number of Corona cases just like Lahore from the beginning, but as shown in the Fig 4. below, the number of cases can be seen an increasing couple of weeks earlier than Lahore. The below chart also shows a spike in the number of recoveries. However, the recoveries also started to jump up from week 20 while the number of fatalities remains steady till week 22 from week 19 on-wards, which is also a much better recovery rate than Lahore. The majority of reasons contributed to this, which will be reviewed later in the discussion section.

Just like any other city, Islamabad also showed a similar

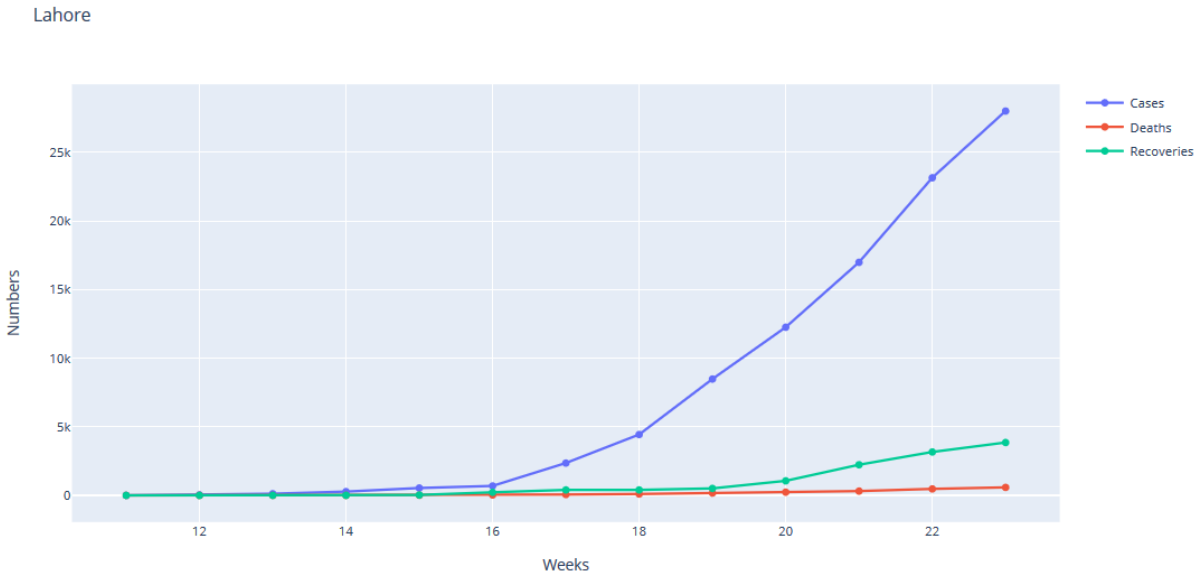

Figure 3. Overall trend in Lahore

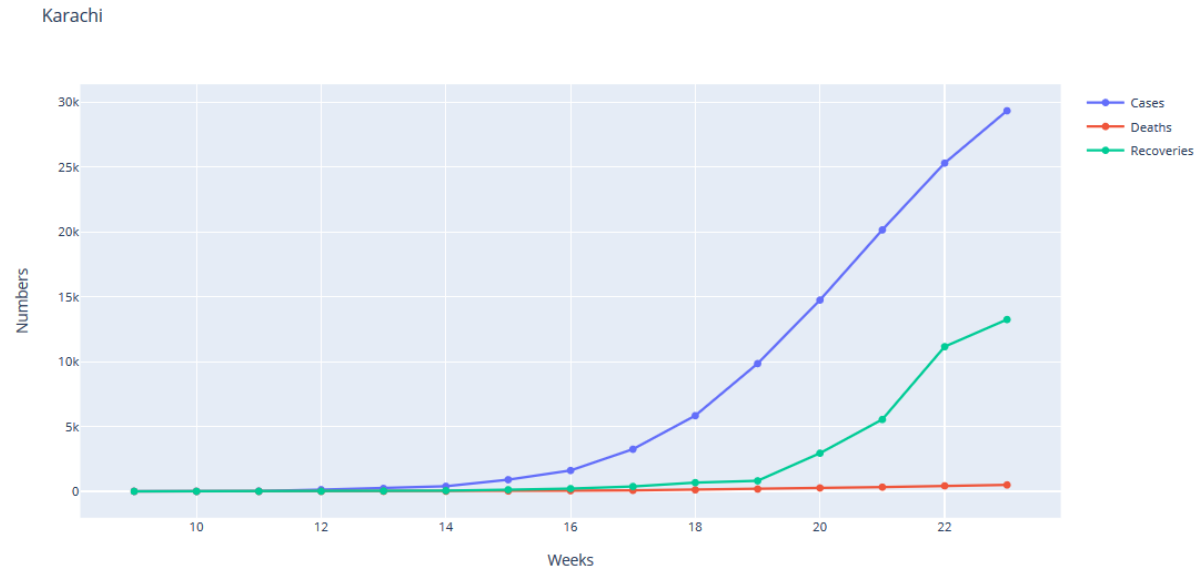

Figure 4. Overall trend in Karachi

trend in the rise of the number of cases as shown in Fig 5. Since the capital has a smaller population than other cities comparatively, the increase in the number of cases at this rate with recoveries increasing after a very long time shows that the inadequate measures were taken in order to control the city's outbreak. Several other factors might be involved in such a fast spread and low recoveries, which is explained later.

Unlike any other city in Pakistan, the number of confirmed cases in Peshawar has grown at a much faster rate, which can also be viewed in the Fig 6. below. But the most considerable number of deaths are also reported in this city, and that can also be seen in the graph. However, the number of deaths and recovery rates are almost parallel.

All the cities in Pakistan have varying temperatures across the year; therefore, it is essential to look at the temperature variable and how it affects the number of COVID-19 cases across the cities.

As shown in the figures 7,8,9,10 below, the number of cases started increasing as the temperature values increased with time.

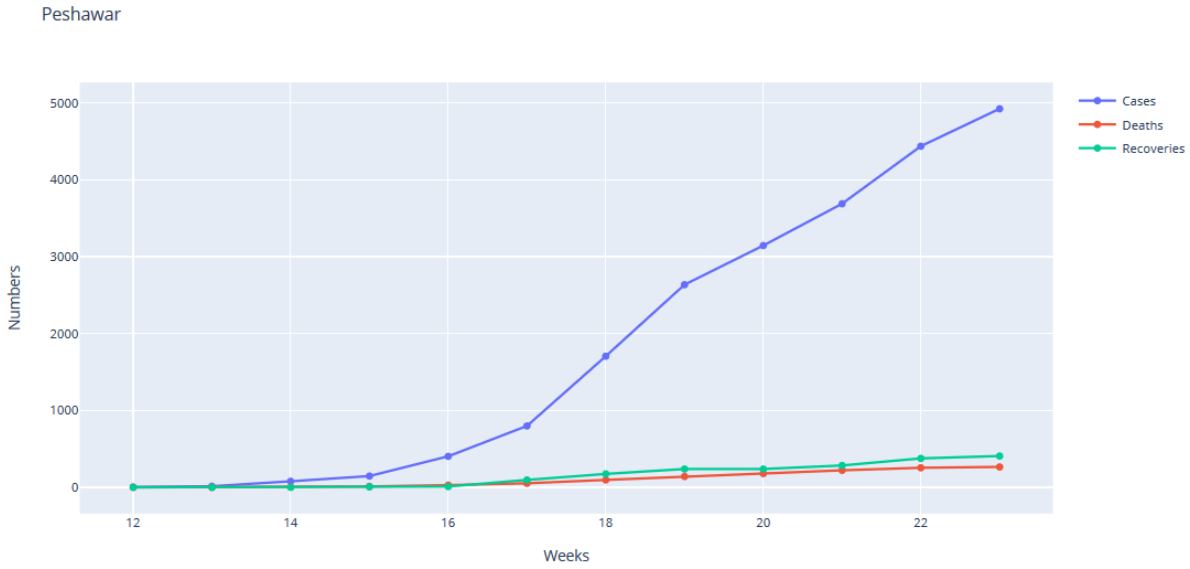

Figure 5. Overall trend in Islamabad

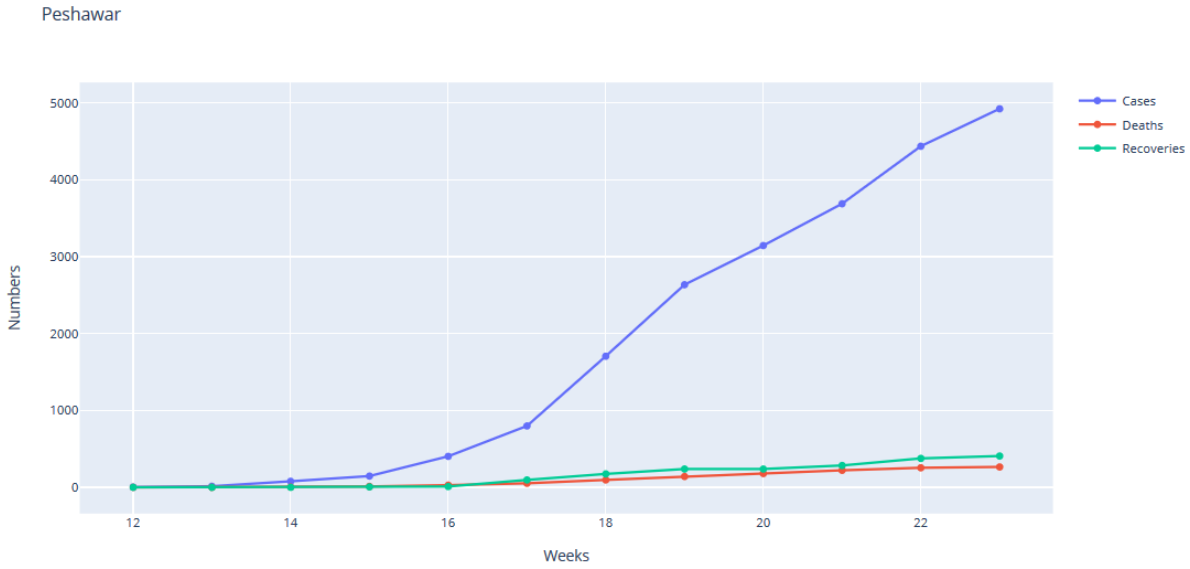

Figure 6. Overall trend in Peshawar

We used the Lasso Model to find out the feature importance of weather components, i.e., Temperature (tempC) and humidity, along with the number of tests performed (tests) due to multicollinearity in the data.

The plots below show that neither temperature no humidity has an impact on our model.

Since the below features (temperature, humidity) are not important for the model. Therefore, it is enough evidence to conclude that the weather components have no impact on the spread of COVID-19 in Pakistan.

Further, by using the O.L.S. regression summary, it can be found out that p-values for the number of COVID-19 tests (polymerise chain reaction) conducted are less than the significance level (0.05). This indicates that the higher the number of tests conducted, the higher number of cases are reported, and they are independent of any weather impact given the data.

Black dots shown in the figure.15 below represents the actual values, while the blue line indicates the predicted values between a 95 percent confidence interval. The red dots in the plot above shows the predicted test data.

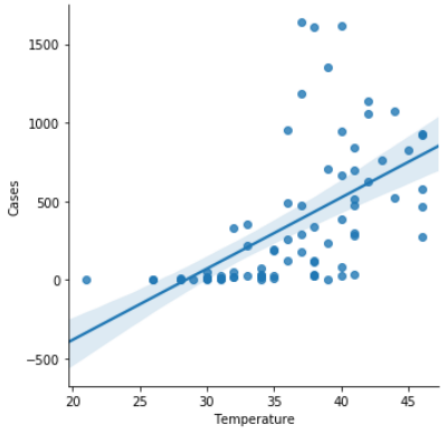

Figure 7. Temperature Wise Cases for Lahore

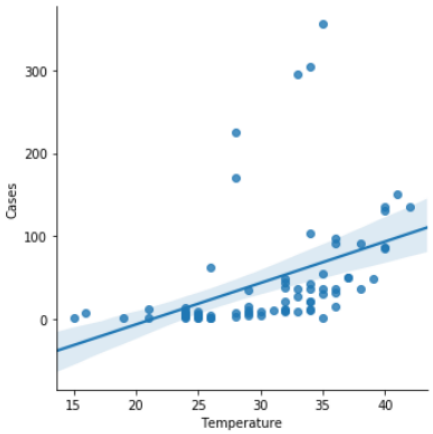

Figure 10. Temperature Wise Cases for Islamabad

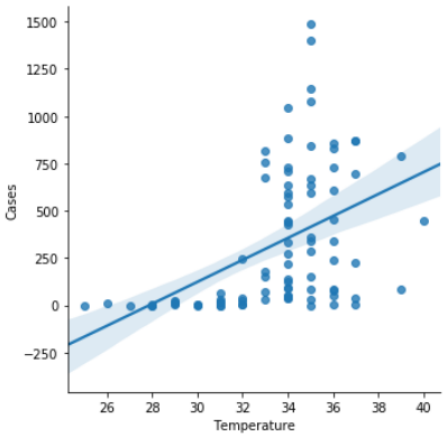

Figure 8. Temperature Wise Cases for Karachi

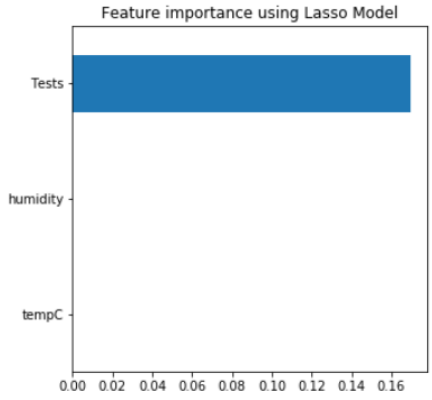

Figure 11. Feature Importance for Lahore

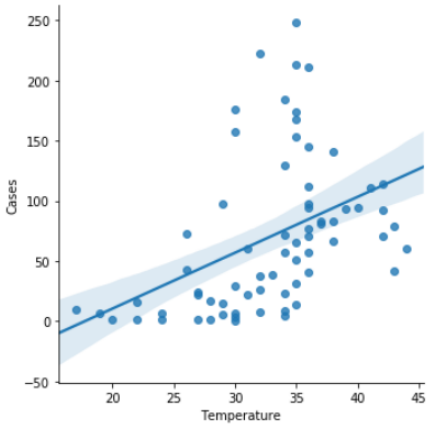

Figure 9. Temperature Wise Cases for Peshawar

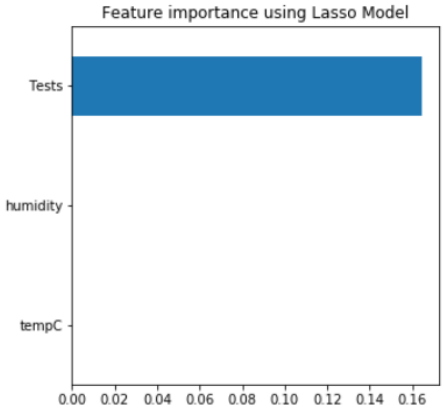

Figure 12. Feature Importance for Karachi

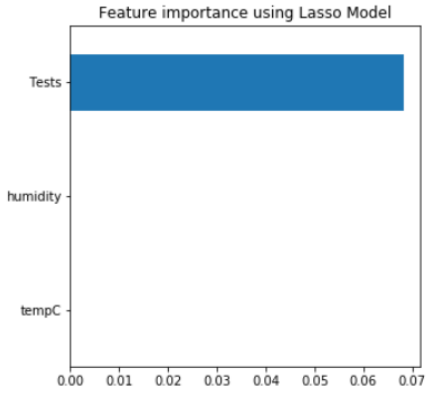

Figure 13. Feature Importance for Peshawar

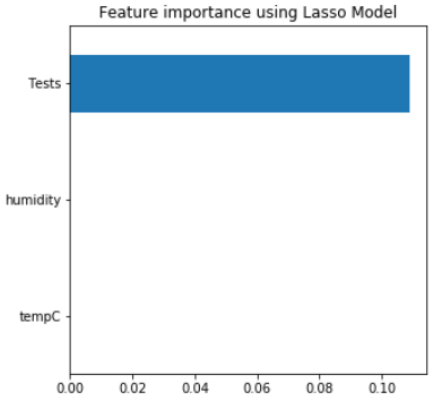

Figure 14. Feature Importance for Islamabad

The mean absolute percentage error (MAPE) of the model is 8.37 percent. It indicates that overall, the points predicted, we are out with an average of 8.37 percent from the actual value.

B. *Forecasting the Impact of COVID-19*

Pakistan followed a province-wise lock down approach. Most of the provinces went under complete lock down starting from March 23, 2020. The lock down from the whole country was lifted on May 9, 2020 [22]. Therefore, the number of days between this period, i.e., 47, is used as a holiday period. A lower window of 2 days is used to include the weekend as March 23 was on Monday, and an upper window of 1 day is used as of May 9, 2020, was a weekend (Saturday), the number of places usually start their operations after the weekend. As our data shows no seasonality, we have not used any seasonal parameters.

Plotting trend component from the model gives us the following graph below shown in figure.16

The below plot shows a linear trend between the number of cases and days without any influence of holidays or seasonality. A curve can be seen at the end of April, which shows the number of cases declined for a shorter period and then began increasing rapidly.

C. *DISCUSSION*

As seen in the results, that weather has no impact on the number of cases in Islamabad, Lahore, Karachi, and Peshawar. Therefore, it is sufficient to claim that there can be many other variables that are responsible for the number of increased cases. One of the key reasons is the lack of social distancing and not following the standard operating procedures (S.O.Ps) announced by the Government of Pakistan and W.H.O. As the population density of Pakistan is 287 sq. Km in 2020 as per United Nations [27], apart from the negligence of the citizens, the social distancing cannot also be maintained sternly, especially in residential neighborhoods of rural areas where the ratio of the size of the house and number of people living in it is small.

The number of confirmed cases also depends on the COVID-19 tests (polymerise chain reaction) performed, as seen from the results. As the number of tests performed earlier was less than now, the number of reported cases was also a few earlier. This is majorly due to a lack of testing facilities readily available to the public in the country. The limited availability of testing kits also slowed down the process of testing. Also, the COVID-19 detection test is relatively expensive if taken privately. Well-off people can only afford it as compared to the majority of the population which cannot afford private health care and relies heavily on government facilities.

Lack of strict lock down policies announced by the government of Pakistan. As per the orders of the Supreme Court, the opening of shopping malls and other non-essential places for the public just before Eid (A joyous religious event for Muslims) resulted in accumulation in the number of cases. Citizens showed extreme negligence and did not follow any S.O.P.s or restrictions during the festive period. As symptoms of the disease take some time to develop, a sudden increase in the number of cases is seen.

As the majority of Pakistan's population lives in rural areas without any proper health care system, the number of cases in major districts surrounded by villages was reported higher in number than the small cities. These people usually have to move to major cities to get themselves treated. Pakistan is a country with not a very good health care system; therefore, the number of beds, medicines, ventilators, and other essential equipment in the hospitals fell short due to sudden pressure resulting in an increased number of deaths and slow recoveries.

5. CONCLUSION

Our study concludes that no impact of the weather or its components significantly affect rising cases in Pakistan's metropolitan cities. Also, the movement of infected individuals to other areas can cause a rapid eruption in those

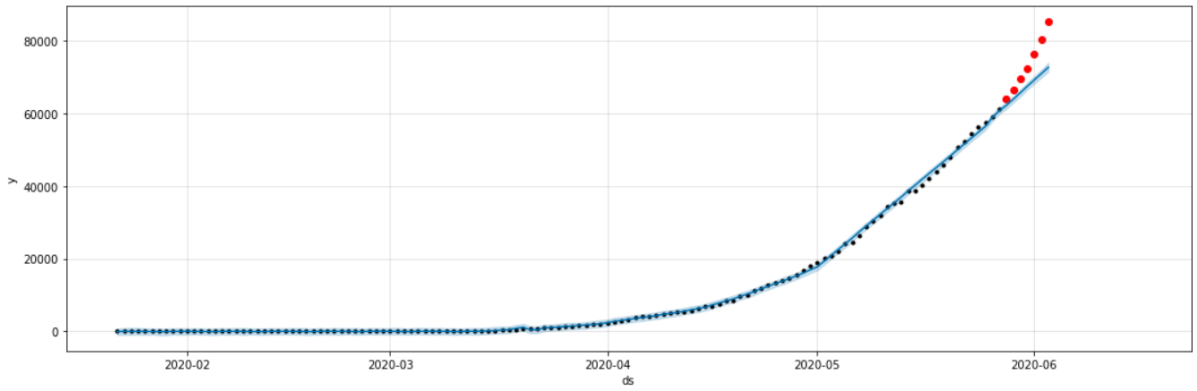

Figure 15. Shows the forecasted values as well as historical data within 95 percent of the confidence interval

1

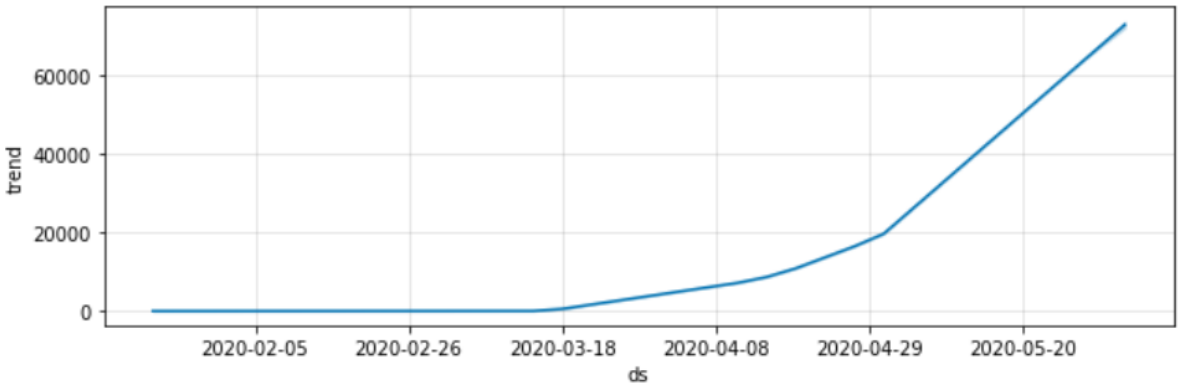

Figure 16. Trend between Dates and number of Cases

areas. No model can determine the movement of people and predict such eruptions. Government instructions for maintaining social distance, isolation, hygiene, and health care can affect the spread and prevent further outbreaks. To date (June 3, 2020), we cannot reach a peak in the number of cases in Pakistan. Therefore, once the peak is obtained, and the number of cases starts declining, we can use different prediction algorithms to identify the trends and produce updated predictions.

References

[1] K. E. Nelson and C. M. Williams, "Early history of infectious disease: Epidemiology and control of," *Infectious disease epidemiology: theory and practice*, vol. 1, 2001.

[2] N. F. Cantor, *In the wake of the plague: the Black Death and the world it made*. Simon and Schuster, 2001.

[3] . World Health Organization *et al.*, "Statement on the second meeting of the international health regulations (2005) emergency committee regarding the outbreak of novel coronavirus (2019-ncov)," 2020.

[4] M. U. Sattar, M. Imran, H. W. Khan, A. Ghaffar, and H. Mushtaq, "Selecting a better classifier using machine learning for covid-19," *International Journal of Computing and Digital System*, 2021.

[5] J. Hopkins, "Johns hopkins coronavirus 15 resource center," *COVID-19 Case Tracker*, 2020.

[6] J. Liu, J. Zhou, J. Yao, X. Zhang, L. Li, X. Xu, X. He, B. Wang, S. Fu, T. Niu *et al.*, "Impact of meteorological factors on the covid-19 transmission: A multi-city study in china," *Science of the total environment*, vol. 726, p. 138513, 2020.

[7] J. Wang, K. Tang, K. Feng, X. Li, W. Lv, K. Chen, and F. Wang, "High temperature and high humidity reduce the transmission of covid-19," *arXiv preprint arXiv:2003.05003*, 2020.

[8] T. Dehesh, H. Mardani-Fard, and P. Dehesh, "Forecasting of covid-19 confirmed cases in different countries with arima models," *MedRxiv*, 2020.

[9] 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*, vol. 724, p. 138226, 2020.

[10] M. A. Cole, R. J. Elliott, and B. Liu, "The impact of the wuhan covid-19 lockdown on air pollution and health: a machine learning and augmented synthetic control approach," *Environmental and Resource Economics*, vol. 76, no. 4, pp. 553–580, 2020.

[11] P. Siriyasatien, S. Chadsuthi, K. Jampachaisri, and K. Kesorn, "Dengue epidemics prediction: a survey of the state-of-the-art based

on data science processes," *IEEE Access*, vol. 6, pp. 53 757–53 795, 2018.

[12] T. M. Chieng, Y. W. Hau, and C. W. Lim, "Ventricular tachyarrhythmias prediction methods and its prognostic features: a review," *International Journal of Computing and Digital Systems*, vol. 8, no. 04, pp. 351–365, 2019.

[13] S. Lakshmi Priyadarsini and M. Suresh, "Factors influencing the epidemiological characteristics of pandemic covid 19: A tism approach," *International Journal of Healthcare Management*, vol. 13, no. 2, pp. 89–98, 2020.

[14] S. Abudalfa and H. Qusa, "Evaluation of semi-supervised clustering and feature selection for human activity recognition," *International Journal of Computing and Digital Systems*, vol. 8, no. 6, 2019.

[15] E. Team, "The epidemiological characteristics of an outbreak of 2019 novel coronavirus diseases (covid-19)—china, 2020," *China CDC weekly*, vol. 2, no. 8, p. 113, 2020.

[16] B. Ahmad, "Intelligent digital twin to make robot learn the assembly process through deep learning," *LGURJCSIT*, vol. 5, no. 3, pp. 65–72.

[17] M. Aftab, M. U. Sattar, M. Shaheryar, H. W. Khan, R. Hassan, S. Shahzadi, and S. Mahmood, "Machine generated deep image captioning with style," *International Journal of Computing and Digital Systems*, vol. 10, pp. 1–15, 2020.

[18] G. Onder, G. Rezza, and S. Brusaferro, "Case-fatality rate and characteristics of patients dying in relation to covid-19 in italy," *Jama*, vol. 323, no. 18, pp. 1775–1776, 2020.

[19] M. U. Sattar, S. Palaniappan, A. Lokman, A. Hassan, N. Shah, and Z. Riaz, "Effects of virtual reality training on medical students' learning motivation and competency," *Pakistan journal of medical sciences*, vol. 35, no. 3, p. 852, 2019.

[20] J. A. Lewnard and N. C. Lo, "Scientific and ethical basis for social-distancing interventions against covid-19," *The Lancet infectious diseases*, vol. 20, no. 6, pp. 631–633, 2020.

[21] Y. Zhang, Y. Zhang, and Z. Liu, "The role of different transportation in the spreading of new pandemic influenza in mainland china," in *2011 19th International Conference on Geoinformatics*. IEEE, 2011, pp. 1–6.

[22] J. S. Jia, X. Lu, Y. Yuan, G. Xu, J. Jia, and N. A. Christakis, "Population flow drives spatio-temporal distribution of covid-19 in china," *Nature*, vol. 582, no. 7812, pp. 389–394, 2020.

[23] M. Naveed, D. Baleanu, M. Rafiq, A. Raza, A. H. Soori, and N. Ahmed, "Dynamical behavior and sensitivity analysis of a delayed coronavirus epidemic model," 2020.

[24] M. Sattar, S. Palaniappan, A. Lokman, N. Shah, U. Khalid, and R. Hasan, "Motivating medical students using virtual reality based education," *International Journal of Emerging Technologies in Learning (iJET)*, vol. 15, no. 2, pp. 160–174, 2020.

[25] P. B. of Statistics, "Blockwise provisional summary results of 6th population & housing census-2017," 2017.

[26] F. K. Khan *et al.*, *A geography of Pakistan: environment, people and economy.* Oxford University Press, 1991.

[27] M. K. Anser, Z. Yousaf, M. A. Khan, A. A. Nassani, M. M. Q. Abro, X. H. Vo, and K. Zaman, "Social and administrative issues related to the covid-19 pandemic in pakistan: better late than never," *Environmental Science and Pollution Research*, vol. 27, no. 27, pp. 34 567–34 573, 2020.

Mehak Narmeen is currently enrolled in M.S. Data Science program at University of Management and Technology, Lahore, Pakistan. At present, she is also working as a Senior Software Engineer in a private software house. She has strong programming skills and also a mentor to her juniors. She has participated and won multiple coding competitions across the country. She has been a UGRAD scholarship holder and spent a semester as an exchange student at Murray State University, KY, U.S.A. Her research interests include Machine Learning, IoT, Enterprise Resource Planning and Data Analysis.

Usman Sattar has more than fifteen years of teaching, training, and consultancy experience in the fields of Information System especially Immersive Technologies, Data Science, and Enterprise Systems System Security. He is PhD (Informatics) – Scholar at Malaysia University of Science and Technology, Malaysia.

Hamza Wazir Khan has seven years of diversified experience in field of Information System and Marketing. Currently, he is working as a Lecturer in Information System Department at University of Management and Technology, Lahore. He has given his services in form of teaching, trainings, and consultancy in fields of Data Science, Digital Marketing, Enterprise Resource Planning, and E Business.

Mehak Fatima is currently enrolled in M.S. Data Science at the University of Management and Technology, Lahore. She is currently working as a freelance data analyst and machine learning engineer. She has participated in various entrepreneurship competitions and also the event organizer at International Conference on Innovative Computing (I.C.) 2 in collaboration with U.M.T. and George Mason University U.S. Her research interests include statistical data analysis, Machine Learning, Computer Vision, Big Data and Natural Language Processing.

Mueen-ud-Din Azad is working as an Assistant Professor of the Department of Economics at the University of Management and Technology, Lahore. He has more than six years of teaching experience. He did M.S. Statistics from the University of California Riverside, USA.His areas of interests include Probability Models, Bio-statistics, Order Random Variables, Statistical modeling, Reliability Analysis and Multivariate Analysis. He has published nine research papers in National and International journals. He has also presented his research at various conferences.

Fatima Ghani has done BS in computer Sciences and currently enrolled in MS Data Science at University of Management and Technology, Lahore. She is also working as Data Analyst at Directorate of General Health Services, Punjab, Lahore. She has a good and sound knowledge of data analysis and reporting of large-scale data specially in health sector. Her research interest includes Deep Learning, Machine learning, Data Analysis and Data Reporting.