International Journal of Innovative Research in Computer Science & Technology (IJIRCST)
ISSN: 2347-5552, Volume-9, Issue-3, May 2021
https://doi.org/10.21276/ijircst.2021.9.3.2
Article ID IRP1147, Pages 11-20
www.ijircst.org

Data Science Approach for Rapid COVID-19 Case Rate Study of Important Climatic Regions in Correlation with Temperature and Humidity Variations of INDIA

Maharishi Kalla[1], Dikshant Bhati[2], Aman Arora[3]

[1] Department of Biotechnology and Bioinformatics, Jaypee University of Information Technology, Solan, H.P, India
[2] Department of Medicine and Surgery, Shimoga Institute of Medical Science, Shivamogga, Karnataka, India
[3] Department of Medicine and Surgery, Jawahar Lal Nehru Medical College, Ajmer, Rajasthan, India

Correspondence should be addressed to Maharishi Kalla; maharishikalla@gmail.com

ABSTRACT- The early pandemic wave came after the imposition of Lockdown by Indian Government. Although it wasn't clear if factors like temperature and humidity have any correlation with the increase of cases in the Locked regions of the state, as studies were on going in virology and biotechnological field for that matter. We took public health vision for the study, by considering five important climatic regions of the nation with different temperature and humidity in a scientific attempt to provide correlation between these factors and rising COVID 19 case counts in India. Crowd source data was used for the analysis in this study. Initial database of confirmed case counts of COVID 19, humidity, maximum and minimum temperature of five important climatic regions was created. The data collected was removed of redundancies and mean average figures were calculated, to perform correlation analysis on the database using SPYDER (The Scientific Python Development Interface). Statistical approaches such as Pearson correlation coefficient and linear regression curve graphical representation was used for the analysis. The analysis significantly presented positive correlation between confirm case count increase and climatic factors (humidity and temperature) in five different climatic regions of India for the period of four lockdowns. The analysis proved the region with greater humidity was suffering on a great frequency with respect to less humid region. Average temperature was nearly equal for every climatic region, nevertheless the analysis presented with figures signifying nearly nil or low positive correlation between confirmed case count ratio and climatic regions with high or low mean average temperature. It was evident that climatic factors had an impact on the transmission and sustaining capability of COVID 19. The analysis of different climatic regions that also during the period of Lockdown is essential to health care sector, government and policy makers while making new policy decisions and taking new measures for prevention of spread of COVID 19 pandemic and any future pandemic scenarios of the level.

KEYWORDS- COVID-19, Climatic factors, correlation, confirmed cases count, India.

I. INTRODUCTION

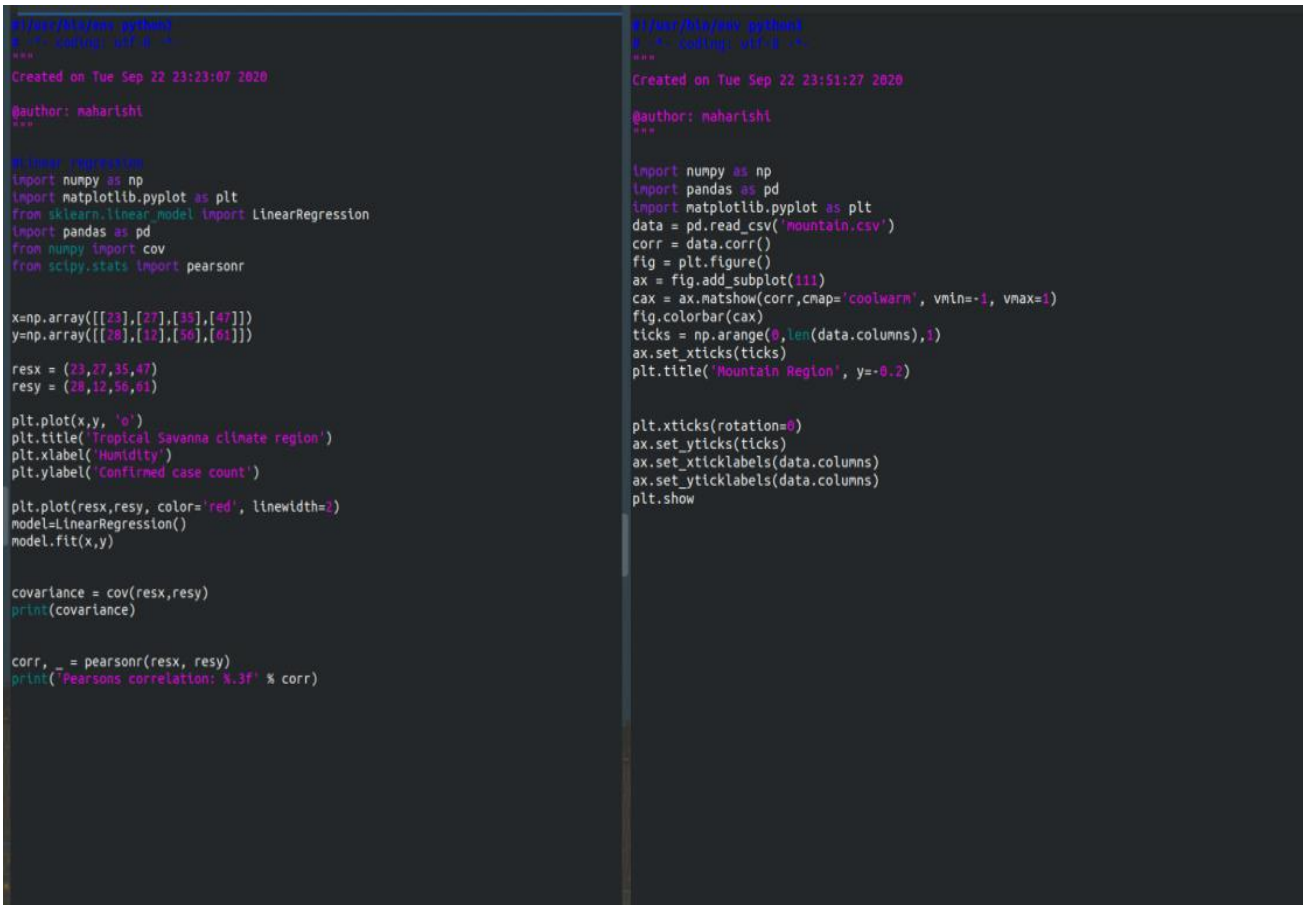
This year humanity has suffered enormously from COVID-19, In INDIA the virus first came into highlights during the end of January of 2020. The first case was noted on 30th January 2020 and since then the growth of infections has increased in rapid manner over short period of time. As of now INDIA has confirmed over 40, 00,000 cases of COVID 19 [1], with this fast immergence of COVID 19 in the country with share of more than 15% of the total cases across the globe. Meanwhile as we surge in the study of SARS-CoV-2 virus itself there is immense number of studies going on for that matter. The government came up with policies and strategies to limit the spread of COVID 19, these policies majorly restricted the movement of people of the nation. This was achieved through Lockdown strategy implied by the government that included of 4 phases in total [2] i.e., Phase 1 (24th march to 24th April), Phase 2 (25th April to 3rd may), Phase 3 (4th may to 17th may), Phase 4 (18th may to 31st may), these strategies gave the government and research agencies some upper hand for the matter of study on the pandemic and the virus but eventually the rate of cases started to increase after partial ending of the Lockdown. It was essential to remove complete lockdown as being a developing country major population income source is occupation that cannot be regulated during Lockdown phase thus Government of INDIA came up with partial Lockdown strategies following full Lockdown phase. With on-going advances in the research community for reliable cure of the viral disease and no such flawless vaccine in hands it is getting difficult

for the individuals and communities to stabilize there life during the pandemic. Whilst we surf for possible ways to study the factors and ratios through data of the cases, total confirmed cases are taken into account for the study. The lockdown phases that government of India issued marked the rise of cases with the regions being separated by another parts of the country fulfilling the objective behind the lockdown strategy itself, this segregations of different types of regions gave way to many studies regarding specific factors for particular regions thus creating a way to search and study about every minute factor responsible for the spread rate of COVID 19. India is a country with many diversity, culture and most importantly climatic conditions. The spread of COVID 19 created alarming issues with government agencies and organizations trying to find the systemic procedure and plan for stabilizing the situation in best manner possible. A related concern was about the virus affectivity in different climatic regions of India. Many of the studies reflected the concern for how factors such as temperature, humidity and fomites involved in the transmission of the virus affect the case rate in many aspects. The human coronavirus associated with the common cold was reported to remain viable only for 3 hours on environmental surfaces after drying, although it remains viable for many days in liquid suspension [3]. These studies showed the importance of such factors which can be taken in account while constructing strategies to minimise the spread of COVID 19 from one region to another. India consists of various important climatic regions spreaded across the country based on different geographic conditions. Most important climatic regions that consist of some major states and the cities are Tropical Savannah region, Tropical and sub-Tropical steppe, tropical semi-arid steppe, Dessert region and Mountain region. Different climatic regions differ in many factors consisting most importantly Humidity and Temperature, for example dessert being the hottest region of the country, costal area under Tropical Arid steppe most humid and Mountains with high altitudes having the least temperature in the country. Study of confirmed cases rates in different climatic conditions will provide us with the results that can out show and differentiate these climatic regions on case severity based on climatic conditions and involvement of climate factors in the sustainability and spread of COVID 19. The Analysis of cases count in correlation with different important climatic regions of the country is important as to know the regions which are more vulnerable to the pandemic and where the virus is more efficiently surviving and transmitted with climate being the major factor. Population of these different climatic regions have biologically and culturally adapted the environment around them, which makes it important to study the case rate of the COVID 19 in these climatic regions communities. The analysis is aimed so as to provide the measure of case rate with correlation to specific population that comes under these major climatic regions. The analysis will help to correlate the spread map of the virus as after lockdown population will start to travel and transmission will increase on a mass level.

II. METHODOLOGY

The cloud source data on COVID 19 cases in India, on which the analysis is based on, is available for download publicly at https://www.covid19india.org/. The portal provides COVID 19 data in cumulative form and daily series, regional data is also published in the same format. The data is consistent and is provided by the Ministry of Health and Welfare. International databases of COVID 19 are also checked for the purpose such as https://coronavirus.jhu.edu/map.html from the Johns Hopkins University and Medicine [4]. The conformed case count for the period of four lockdowns was recorded, based on four specific cities covered under the five important climatic regions of the country. The data of these regions was accumulated and mean average was calculated for each Phase of the lockdown in the region. Gathering data for every city coming under following climatic regions was not feasible reason being data from some regions and districts which were very small and backward was not available. For this reason four important cities in such a manner that every area of these climatic regions are covered for the purpose of calculating the mean average of conformed case count. The rate of testing differs from region to region but the main focus here was to accumulated data of those cities which have fluent data flow with good testing while being the part of these climatic region. The cities that were taken under following five important climatic regions: Chhattisgarh, Nagpur, Chandrapur, Balaghat these four major cities fall under Tropical Savannah climate region. Mau, Azamgarh, Patna, Jharkhand falling under Tropical sub-Tropical steppe. Chennai, Vishakhapatnam, Mysore, Tiruvananthpuram falling under Tropical semi-arid steppe. Jaisalmer, Barmer, Bhuj, Palanpur falling under Dessert region. Kullu, Lahaul, Leh, Kinnaur falling under Mountain region of the country. Temperature and humidity data of climatic region was available on Indian Meteorological department, Ministry of Earth Sciences [5] https://mausam.imd.gov.in/, other commercial databases were also used for collection of data with precise resource and valid data such as accuweather https://www.accuweather.com/en/in/national/satellite [6] and weather and climate from Amsterdam https://weather-and-climate.com/ [7] the data was recorded for the same specific cities used in the data collection of conformed cases, to make the data correct for the study of correlation between three variable data sets. After gathering the data first step included making data redundant free and less complicated with same accuracy and potency of the units, it was achieved by calculating mean average data from all phases in the dataset for example: the maximum and minimum temperature data of Tropical Savannah climatic region consisted four cities. The data of four phases of lockdown was obtained of these cities consisting of Minimum and Maximum Temperature during the time of four Lockdown phases i.e., Phase 1 (24^{th} march to 24^{th} April), Phase 2 (25^{th} April to 3^{rd} may), Phase 3 (4^{th} may to 17^{th} may), Phase 4 (18^{th} may to 31^{st} may). This made calculation of mean average of Phase 1 and all the following phases of the climatic region possible by adding

all the data of Temperature from Phase 1 of different four cities under Tropical Savannah and dividing it. With four being the total number of variables, with this approach mean data was calculated of every phase under five important climatic regions. Mean average formulae: $\frac{\sum \text{of the terms}}{\text{no of total terms}}$,the mean average data of all three variables i.e., Minimum maximum Temperature, Humidity and Conformed case count made a viable data for using regression correlation function on the same. For the purpose of studying the correlation between COVID 19 confirm caste rate rise and Temperature and Humidity to find out if climatic condition may possibly affect the transmission and rise of COVID 19 case rate, some correlation function were used namely Pearson correlation, Linear Regression graph and multiple correlation graph (for case with multiple variables i.e., min max temperature in correlation with conformed case rate). Pearson Correlation coefficient is used for the purpose of comparing the variables such as Humidity and Conformed case fluctuation in the climatic regions as it determines the measure of strength of linear association between two variables datasets. It attempts to draw a line of best fit between the data of two variables, while the Pearson correlation coefficient indicates how far away all these data points are to this line of best fit i.e., how well the data points fit in the model [8]. The Scientific Python Development Interface (SPYDER) was used to find the Pearson correlation coefficient and linear regression graph as well as multiple correlation as there is a significant convergence of interests in the research community efforts to advance the development and application of software resources (capable of handling the relevant mathematical algorithms to provide scalable information) for solving data science problems. We used Anaconda as anaconda is one of the many open source platforms that facilitate the use of open source programming languages (SPYDER, r) for large-scale data processing, predictive analytics, and scientific computing [9]. Using these algorithms and statistical function we obtained the linear regression curve of phases between humidity and conformed case count of the following five climatic regions also we obtained the multiple regression graph to estimate the effect of temperate on the conformed case count in these climatic regions during the period of lockdown. This cure and Pearson correlation aims to tell us about the severity of the increase case count in different climatic regions of India. The results from different climatic regions based on the data was compared so as to find out which region climate was most suited for the transmission and sustainability of the virus as we compared case count with the results and the curve obtained.



```
#!/usr/bin/env python3
# -*- coding: utf-8 -*-
"""
Created on Tue Sep 22 23:23:07 2020

@author: maharishi
"""

#Linear regression
import numpy as np
import matplotlib.pyplot as plt
from sklearn.linear_model import LinearRegression
import pandas as pd
from numpy import cov
from scipy.stats import pearsonr


x=np.array([[23],[27],[35],[47]])
y=np.array([[28],[12],[56],[61]])

resx = (23,27,35,47)
resy = (28,12,56,61)

plt.plot(x,y, 'o')
plt.title('Tropical Savanna climate region')
plt.xlabel('Humidity')
plt.ylabel('Confirmed case count')

plt.plot(resx,resy, color='red', linewidth=2)
model=LinearRegression()
model.fit(x,y)


covariance = cov(resx,resy)
print(covariance)


corr, _ = pearsonr(resx, resy)
print('Pearsons correlation: %.3f' % corr)
```

```
#!/usr/bin/env python3
# -*- coding: utf-8 -*-
"""
Created on Tue Sep 22 23:51:27 2020

@author: maharishi
"""

import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
data = pd.read_csv('mountain.csv')
corr = data.corr()
fig = plt.figure()
ax = fig.add_subplot(111)
cax = ax.matshow(corr,cmap='coolwarm', vmin=-1, vmax=1)
fig.colorbar(cax)
ticks = np.arange(0,len(data.columns),1)
ax.set_xticks(ticks)
plt.title('Mountain Region', y=-0.2)


plt.xticks(rotation=0)
ax.set_yticks(ticks)
ax.set_xticklabels(data.columns)
ax.set_yticklabels(data.columns)
plt.show
```

Fig. 1: Python script for finding correlation

III. RESULTS

A. Confirmed Case Count

The data for five important climatic region climatic region consisting the average conformed case count during the period of four Lockdown Phases i.e., Phase 1 (24th march to 24th April), Phase 2 (25th April to 3rd may), Phase 3 (4th may to 17th may), Phase 4 (18th may to 31st may), showed rise on every curve. Since 30th March India had reported more than 112,000 cases of COVID nationwide hence beginning the lockdown phase from 24th of March. **Table 1** presents the mean average data of conformed case count from following five important climatic region of India. The data represents increase in the following states of India falling under these climatic regions with different pace and signs making it vital to correlate them with climatic factors of the region. The conformed case count of the population in Table 1 represents the native population of the region that is people biologically and mentally adapted to the climatic factors of following five important climatic regions of India. With this the food intake the living conditions and various other aspects are changed from region to region this is necessary for the study so as this population was locked in their native climatic region making the study of factors such as humidity and temperature on the virus transmission and sustainability precise.

Table 1: Average confirmed case count

Climatic region	Phase 1	Phase 2	Phase 3	Phase 4
Tropical savannah region	28	12	56	61
Tropical and sub Tropical steppe	11	3	31	49
Tropical semi Arid steppe	163	236	1336	2032
Dessert region	47	3	25	44
Mountain region	4	5	0	2

Source: authors based on https://www.covid19india.org/

[a]note: mean average is taken from four important cities falling under following climatic regions of the country

B. Average Temperature data of Climatic Regions

Temperature of every region consists of daily Maximum and Minimum reading in Celsius or Fahrenheit. The temperature data in this study is Celsius. The temperature among different climatic regions of the country varies in different scales. For the period of lockdown phases in the country, the readings of temperature showed hotter climate in most of the regions as all the lockdown nearly came in the summer period of the country. The data however was important so without taking any exceptions or consideration the data obtained was pure and precise.

Table 2 presents the average data used for studying the correlation with respect to the Case count of the climatic regions in the period of four Lockdown Phases i.e., Phase 1 (24th march to 24th April), Phase 2 (25th April to 3rd may), Phase 3 (4th may to 17th may), Phase 4 (18th may to 31st may). It was recorded the average temperature of mostly every climatic region was ~40/26 degree Celsius making the environment of hot nature reason behind this result was the summer months in which the lockdown phases were implemented. The results denoted that only hot climate study was possible for finding out the correlation between the temperature and mean average confirmed case count of the climatic region. Making the average minimum temperature data of less importance none the less correlation was performed with both Minimum and Maximum temperature to gain a clear sight of the correlation factor study.

C. Average Humidity Data of Climatic Regions

Humidity or the concentration of water vapour present in air is vital for any study relate to climate. Relative humidity data was chosen for the study it represents humidity in percentage. Relative humidity readings vary from region to region in the country with greater difference that that of temperature making it viable factor to consider while studying the viral factors in correlation with the climatic regions of the country. **Table 3** presents the mean average humidity data of following five important climatic regions of India in the period of four Lockdown Phases i.e., Phase 1 (24th march to 24th April), Phase 2 (25th April to 3rd may), Phase 3 (4th may to 17th may), Phase 4 (18th may to 31st may). The data indicated Tropical semi-Arid steppe

climatic region being the most humid zone of all the five most important climatic regions of India. Various researches have suggested the effect of Humidity on SARS-CoV-2 virus responsible for the Pandemic. This result will pin point out the severity of the Humidity factor in relation to Public Health besides from a Virological study point of view, developing Public Health knowledge on the matter.

Table 2: Mean average temperature data

Climatic region	Phase 1	Phase 2	Phase 3	Phase 4
Tropical savannah region	40/23	42/23	41/23	44/28
Tropical and sub-Tropical steppe	40/23	41/25	41/25	44/28
Tropical semi-Arid steppe	36/26	36/27	36/27	37/28
Dessert region	42/25	43/27	44/25	45/28
Mountain region	26/13	30/16	30/16	34/17

Source: authors based on https://mausam.imd.gov.in/ and https://www.accuweather.com/en/in/national/satellite.
[a]note: mean average is taken from four important cities falling under following climatic regions of the country.
[b]note: "x/y" x being the highest average temperature recorded for a phase, y being the lowest temperature recorded for a phase, with degree Celsius being the unit of temperature.

Table 3: Mean average Humidity data

Climatic region	Phase 1	Phase 2	Phase 3	Phase 4
Tropical savannah region	23	27	35	47
Tropical and sub-Tropical steppe	40	50	60	62
Tropical semi-Arid steppe	71	70	71	72
Dessert region	29	26	28	33
Mountain region	52	44	35	35

Source: authors based on https://mausam.imd.gov.in/ and https://weather-and-climate.com/
[a]note: mean average is taken from four important cities falling under following climatic regions of the country.
[b]note: the readings denote relative humidity in percentage figures

D. Linear regression curve and correlation between Humidity and Rising Confirmed case count of COVID 19 in Climatic regions of India

Correlation or relation between two variables which gives the direction relationship between the variable. COVID 19 spread in the world urged to study all the factors affecting the spread of the virus. Statistical approach is a viable method to find and study the factors with respect to COVID 19 pandemic. In this study we found the Correlation coefficient and linear regression curve between Humidity and Confirmed case count of following important climatic regions of India. Use of python for the same made the calculation more precise, SPYDER the synthetic python development interface was commanded with the calculative

algorithms. The results obtained from the interface presented certain outcomes for every climatic region. Important point of the study i.e., the data and study revolves around the four Lockdown phases thus every point on the graph and every result is obtained and studies with same approach in mind. Figure 1 presents the outcome of Linea regression curve obtained from scientific python development interface. The result presented is what was expected of it with linear relationship between both the independent variables i.e., conformed case count and humidity of the region increasing with each other. That was a pre assumed result as the total no of confirmed case count was increasing on daily basis, although the aim to find the relationship between the climatic region factors such as humidity was the main aspect of the study. For the same it is totally noticeable that the region with high humidity is suffering from COVID 19 on much greater scale than other climatic regions. The climatic region of Tropical semi-Arid steppe (south costal area of India) shares nearly majority of the average conformed case count of all the five climatic regions taken in the study (during the period of Four Lockdowns) i.e., 90%. It is also clear the region of the Tropical semi-Arid steppe was the most humid region of all the five important climatic regions of India with average 71 % relative Humidity in the region. This suggested the nature of virus during the period of Lockdown as it is most affective in getting transmitted and surviving in the humid climate of south coastal regions of India. The linear regression line for the case of Tropical semi-Arid steppe came out to be uniform in progression. Whereas the least humid region The Dessert region of the country shared very few of the confirmed case count although it was clear even in the Dessert region the Conformed case count increased with the increase of Humidity at quick pace. As for other regions the linear regression result was no different same pattern followed as rest of the regions had almost same average Relative Humidity throughout the lockdown, except Tropical semi Arid steppe region. Further Pearson correlation coefficient (r) presented the direction of the correlation it was clear from the graphical presentation that relation was going in positive direction, still every climatic region presented with different Pearson correlation coefficient. Table 4 depicts obtained Pearson correlation coefficient for every climatic region. Obtained Pearson correlation coefficient, presented a clear picture with respect to humidity. Although every region had a strong positive relationship with increasing case count as depicted by the Linear regression curve tropical savannah region and the region of Tropical semi-Arid steppe had the strongest of the relationship with r value of 0.832 and 0.812 respectively.

Table 4: Pearson Correlation Coefficient

Climatic Regions of INDIA	**r**
Tropical savannah region	0.832
Tropical sub-Tropical steppe	0.805
Tropical semi-Arid steppe	0.812
Dessert region	0.791
Mountain region	0.762

E. Correlation between Temperature factor and Rising Confirmed case count of COVID 19 in Climatic regions of India

Correlation between Temperature (including both maximum and minimum) and confirm case count was obtained using different approach in SPYDER as this data had three independent variables in total. By using multiple correlation algorithms to find correlation using multidimensional dataset, graphical result was obtained. Through this approach three different datasets can be visualized as graphical representation of correlation which makes the interpretation of the so obtained result understandable to individuals with non-statistical background as well. Figure 2 presents the multi correlation between max min Temperature and Confirmed case count of the following climatic regions of India, obtained from SPYDER python interface. The colour of the squares represents the correlation between three data sets. The dark red colour represents perfect positive correlation while the grey represents no correlation and dark blue represents perfect negative correlation. Which can be seen in the Figure 1 as well as the correlation between two same variables is depicted in dark red. However in this study, The results depicted maximum temperature had the edge with

getting positive correlation with confirmed case count in three of the following climatic region namely Tropical and sub-Tropical steppe, Tropical semi-Arid steppe and Tropical Savannah climate region. While increasing temperature effect can be seen clearly noticeable result was obtained from the Mountain region and Dessert region of the country. Dessert region confirmed case count nearly had no correlation with the maximum temperature while Mountain region showed negative correlation with both the temperature. This result can be a result of the geographical condition of the region with migration of people from these regions being nearly very less, as these regions mainly had population of tribes and communities which comparatively migrate less from the population of rest three climatic regions. Considering the role of Humidity and temperature it was evident that different climatic regions correlate with humidity and temperature on a significant level, the study is relevant to its nature of assessment. The result obtained presents the sketch of how climatic conditions can affect pandemic in daily life of native population residing in different climatic regions of India.

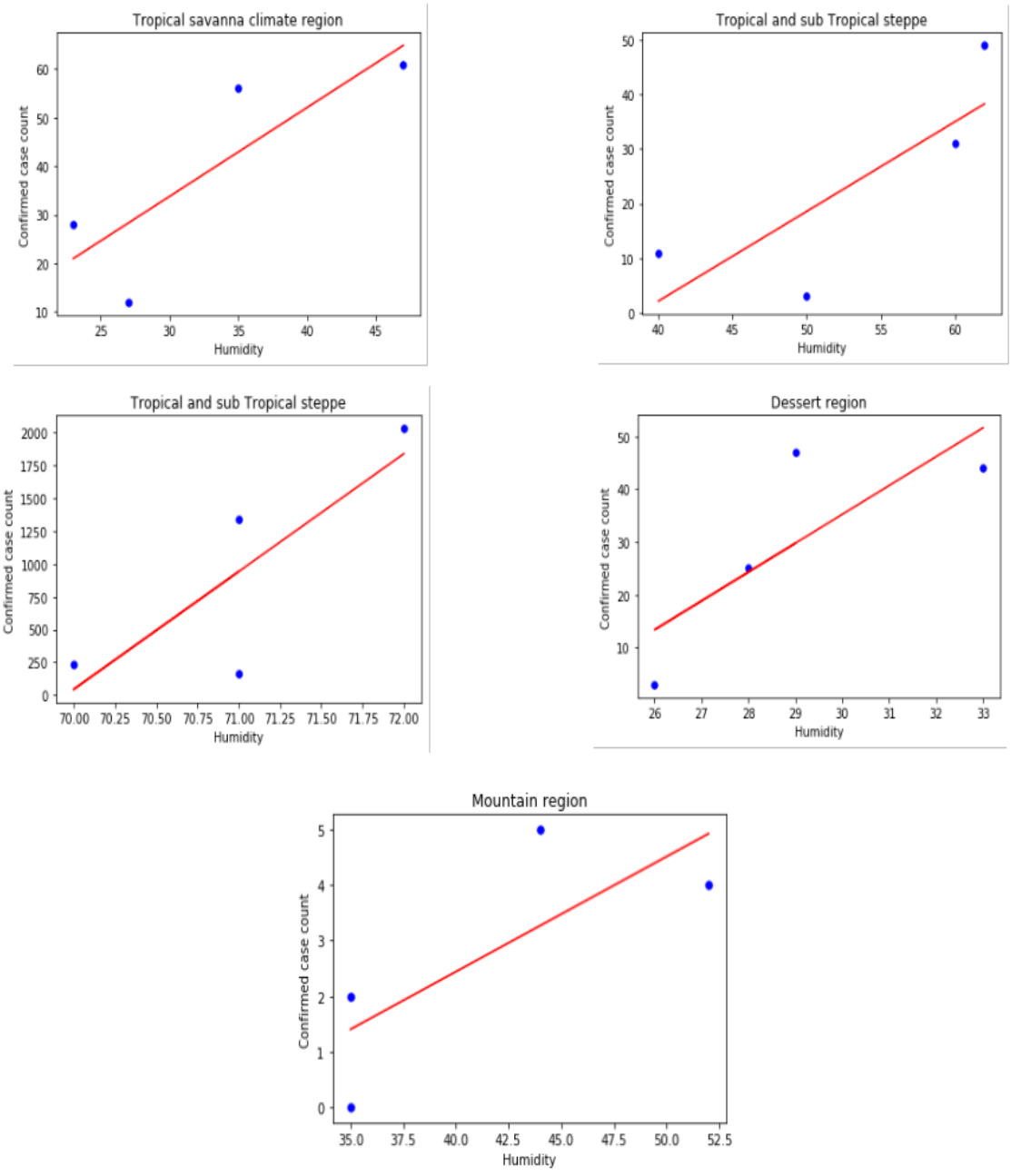


Fig. 2: Linear regression curve between Humidity and confirmed case count of five important climatic region of India

In Figure 2, four dots represent the four lockdown phases i.e., Phase 1 (24th march to 24th April), Phase 2 (25th April to 3rd may), Phase 3 (4th may to 17th may), Phase 4 (18th may to 31st may), While red line represents linear regression line for lockdown period between humidity and confirmed case count.

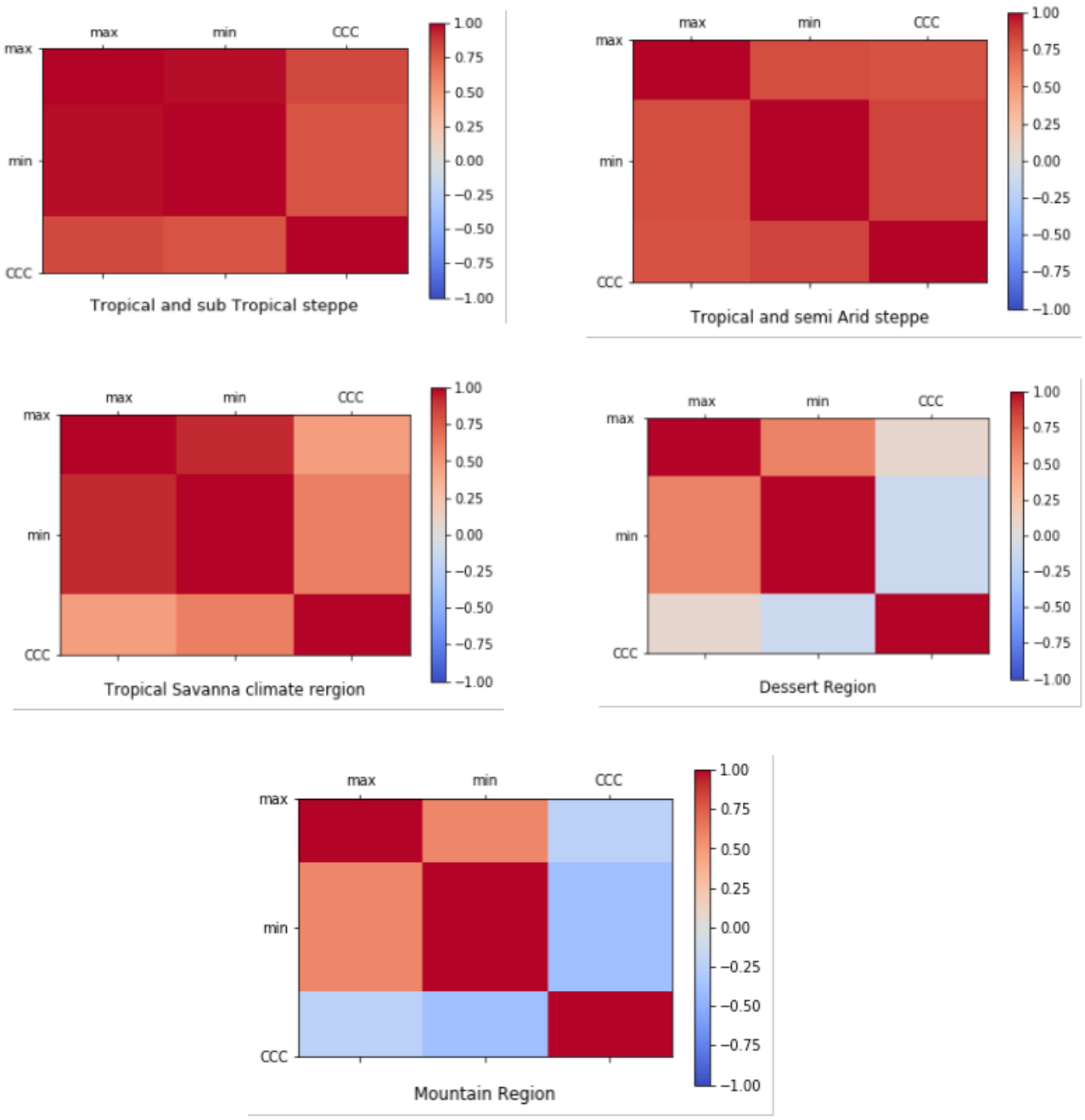


Fig. 3: Presents the multi correlation between max min Temperature and Confirmed case count of the following climatic regions of India

In Figure 3, "CCC" denotes Average conformed case count, "max" denotes average maximum temperature, "min" denotes average minimum temperature of the following climatic region, and the colour bar in right depicts the correlation between the factors.

IV. DISCUSSION

The data analysis presented following salient inferences, firstly different climatic regions of India had different correlation with the temperature and humidity factors. As every factor is important when we study and analyze about pandemic climatic conditions study always becomes

necessary asset to the research. Although many factors are responsible for the transmission and sustainability of the virus which shares equal amount of importance as well. Secondly correlation between Humidity and Confirmed case count significantly depicted the climatic region i.e., Tropical semi-Arid steppe region geographically located near the costal lines of south India was the hot spot for the pandemic. The case ratio was so immense with respect to other four climatic region of India. It is essential to take in the consideration of the living conditions of the native population of this region which is formed in way to live in this humid region of India, as it gives the study more angles to find more better ways to research on community level to slow the pace of virus and keep the virus transmission at lowest. Thirdly the temperature correlation with conformed case count was important to the study as during the time period of four lockdowns the country was in summer season phase with average temperature of every region hotter and equal with minute differences except the mountain region due the altitude difference. To the astonishment of the results obtained the correlation of temperature with confirmed case count varied majorly with all the four important climatic region of the country.

A. Limitations of Study

The data analysis study still had certain limitations, Firstly the important climatic regions take in account for the study had nearly the total area of ~50,000 km^2 to 1,00,000 km^2. Data retrieval for this much area was not possible in terms of availability as data for only major cities and some towns was available. Also any false data would have shown redundancy in the database thus corrupting the analysis and the results. This particular limitation is very common in developing countries, making the study little bit edging towards the region of more population that is urban regions. The second limitation includes confirmed case count, as the study was focused on the four lockdown phases. The data of confirmed case count which was recorded before the period of first Lockdown was also taken in consideration although the frequency was negligible in four of the five climatic regions the Tropical and semi-Arid steppe was an exception. This exception in Tropical and semi-Arid steppe climatic region was due to the high amount of the confirmed case count before the start of the lockdown phase one itself, as exception occurred only in one of the five climatic regions it was included in the data thus affecting the correlation between the factors included in the study. The third limitation came with the multiple variables in the analysis of maximum temperature, minimum temperature and confirmed case count ratio, as for the linear regression line was not included in the analysis for the following three variables rather a different approach was used which was different from the method of analysis used for correlating humidity and confirmed case count.

V. CONCLUSIONS

In context to the data analysis study of temperature and humidity correlation with COVID 19 it presented the fact that these factors are in direct correlation with the intensity of the spread in various regions of India that to with different frequency. The knowledge of data analysis can become boon to the sector of public health. Greater the knowledge greater will be the chances of defending humanity against this and the coming pandemics. It is significant to study and find every possible manner to assist the policy makers and public health organizations with every aspect of the pandemic.

CONFLICTS OF INTEREST

The authors declare that they have no conflicts of interest.

REFERENCES

[1] Government of India COVID-19. Statistical information system. https://www.covid19india.org/. Updated 2020. Accessed September 1, 2020

[2] Ministry of Health and Family Welfare. Updates on COVID-19. https://pib.gov.in/PressReleasePage.aspx?PRID=1615049. Updated 2020. Accessed May 23, 2020

[3] J. S. Malik Peiris, S. Y. Lam, L. L. M. Poon, K. Y. Yuen, and W. H. Seto (2011) The Effects of Temperature and Relative Humidity on the Viability of the SARS Coronavirus. Hindawi 2011;2011:734690.

[4] COVID-19 Dashboard by the Center for Systems Science and Engineering (CSSE) at Johns Hopkins University. https://coronavirus.jhu.edu/map.html. Updated 2020. Accessed September 5, 2020.

[5] Meteorological department, Ministry of Earth Sciences. Climatic Information system.https://mausam.imd.gov.in/. Updated 2020. Accessed September 2, 2020

[6] Accuweather, commercial weather forecast information organisation. https://www.accuweather.com/en/in/national/satellite. Updated 2020. Accessed September 5,2020.

[7] Weather and Climate Amsterdam, commercial weather forecast information organisation. https://weather-and-climate.com/average-monthly-Rainfall-Temperature-Sunshine-in-India. Updated 2020, Accessed September 5, 2020.

[8] Statistical guide. https://statistics.laerd.com/statistical-guides/pearson-correlation-coefficient-statistical-guide.php. Updated 2020, Accessed September 15, 2020.

[9] Akhil Kadiyala, Ashok Kumar (2017). Applications of Python to evaluate environmental data science problems. AIChE 2017;36(6):1580-1586

ABOUT THE AUTHORS

Mr. Maharishi Kalla Under Graduate student of Department of Biotechnology and Bioinformatics, Jaypee University of Information Technology, Solan, H.P, India. His main area of research and work is in field of Biological Data Science, Bioinformatics and

Computational Biology.

Mr. Dikshant Bhati, Under Graduate student, Department of Medicine and Surgery (MBBS), SHIMOGA Institute of Medical Sciences, Shivamoga, Karnataka, India. His area of research and interest is in field of community medicine.

Mr. Aman Arora, Under Graduate student, Department of Medicine and Surgery (MBBS), Jawahar Lal Nehru Medical College, Ajmer, Rajasthan, India. His area of research and interest is in field of pharmacology.