Spatiotemporal Analysis of Weather Effects on COVID-19 Pandemic Transmissions in Select US Counties

Zhenkun Tian1, *, Chuixiang Yi2,3, Yingying Fu4, Jacqueline Singer2, Qin Zhang5

1Department of Mathematics and Computer, China University of Labor Relations, Beijing, China
2School of Earth and Environmental Sciences, Queens College, City University of New York, New York, USA
3Earth and Environmental Sciences Department, Graduate Center, City University of New York, New York, USA
4School of Mathematics and Statistics, Beijing Technology and Business University, Beijing, China
5Institution of Water and Environment Research, Dalian University of Technology, Dalian, China

Email address:
tzk@culr.edu.cn (Zhenkun Tian), chuixiang.yi@qc.cuny.edu (Chuixiang Yi), fiuyath@btbu.edu.cn (Yingying Fu), jacqueline.singer@baruchmail.cuny.edu (J. Singer), zhangqinhan@mail.dlut.edu.cn (Qin Zhang)
*Corresponding author

To cite this article:

Received: September 6, 2020; Accepted: September 29, 2020; Published: October 17, 2020

Abstract: The COVID-19 pandemic has become a global crisis with enormous uncertainty. This study aims to explore what role weather plays in pandemic transmission. We hypothesize that weather conditions (temperature, wind speed, and precipitation) significantly influence the transmissibility of the disease and the number of infected people. We tested this hypothesis by analyzing weather variable in moving-average-windows that varied from 1 to 30 days, and daily new confirmed cases observed from 23 counties in the United States, during the period of January 22 to August 19, 2020. We found consistent results that the moving average temperature over 10 days (Tavg10), the moving average wind speed and the moving average amount of precipitation over 28 days (Wavg28, Pavg28) were the meteorological parameters most closely linked to the outbreak and growth of new cases of COVID-19 in the US. The correlation statistics differed regionally: (1) temperature is negatively correlated to the outbreak of COVID-19 in the Northeastern US and positively in other areas; (2) wind speed is negatively correlated to the COVID-19 pandemic in the Southeastern US while positively in other areas; and (3) precipitation holds a positive correlation on the east coast of the US and a negative one on the west coast. Our results suggest that meteorological factors may play a significant role in COVID-19 pandemic transmission in the US and should be considered by policy makers and crisis administrators.

Keywords: COVID-19, Lag Effect, Temperature, Wind Speed, Precipitation, Pandemic

1. Introduction

In December 2019, a total of 41 cases of pneumonia of unknown etiology which is a new type coronavirus were confirmed in Wuhan city, Hubei Province, China [1]. The World Health Organization (WHO) named the disease as COVID-19 (coronavirus disease 2019) after it spread to several countries outside of China in the situation report-22, and the COVID-19 outbreak was declared to constitute a Public Health Emergency of International Concern [2]. The COVID-19 is transmitted from human to human by respiratory droplets, with the symptoms of fever, cough and shortness of breath [3, 4]. The coronavirus infection clustered within groups of humans in close contact, and death can occur in severe cases [5, 6].

According to the data from the Center for Systems Science and Engineering at Johns Hopkins University (JHU CSSE), more than 5 million COVID-19 cases have been confirmed in the US and more than 173 thousand cases of mortality occurred since the first case reported in Washington State on January 22, 2020 [7]. The US has been the epicenter of the world [8], and the rapid spread of COVID-19 led to the
closure of various industries such as factory, schools, competition league, and so on. As of August 19, 2020, California, Florida, New York, Texas and New Jersey are the top 5 states with confirmed cases in US, and the number of confirmed cases is still rising in almost all the states [7].

In the past 2 decades, the occurrences of Severe Acute Respiratory Syndrome (SARS) and Middle East Respiratory Syndrome (MERS) were associated with climatic factors according to previous studies [9, 10]. In particular, temperature plays an important role in both laboratory and natural conditions [9, 11]. The climatic conditions can be considered as good predictors of coronavirus disease such as SARS, MERS, and COVID-19. Marcos et al. (2020) studied the association between climate variables and global transmission of SARS-CoV-2 and identified a negative correlation between confirmed coronavirus cases and the average temperature by country [12]. Peng et al. (2020) concluded that temperature was an environmental driver of the COVID-19 outbreak in China and temperatures above 8 to 10 Celsius degree were associated with decreased COVID-19 daily confirmed cases rate [13]. Many studies that focused on correlation between weather and COVID-19 pandemic in other locations of the world have been conducted and drew similar conclusions [14-19].

Unlike previous studies, the purpose of this study is to analyze the correlation between weather conditions’ lag effect and COVID-19 pandemic. Besides, the spatial distribution of correlation in different climate zone is also be analyzed. Given that the incubation period of COVID-19 varies from 1 day to 14 days, moving averages of the meteorological parameters with window from 1 to 30 days were considered in this paper.

2. Materials and Methods

2.1. Study Area

Study area was chosen from the worst-hit counties of 23 states of the conterminous United States with severe outbreaks. Figure 1 shows the COVID-19 pandemic in and the locations of the 23 counties by August 19, 2020.

![Figure 1. COVID-19 outbreak in study area of 23 counties and US states from January 22 to August 19, 2020. Data accessed from JHU CSSE, 2020.](https://github.com/CSSEGISandData/COVID-19)

2.2. Data Collection

The COVID-19 data for the period from January 22 to August 19, 2020 were downloaded from JHU CSSE GitHub website which collects the global total COVID-19 cases and updates daily from official source such as WHO, European Centre for Disease Prevention and Control (ECDC), US Centers for Disease Control and Prevention (US CDC), etc. [7]. Table 1 lists the number of total COVID-19 cases, days to 100th case, and date of the first case of select counties.
The weather data for the same period were obtained from the National Oceanic and Atmosphere Administration database [20]. The daily data from meteorological stations consists of temperature average $T$ (°C), temperature maximum $T_{\text{MAX}}$ (°C), temperature minimum $T_{\text{MIN}}$ (°C), wind speed $W$ (m/s) and precipitation $P$ (mm).

In order to investigate the possible time lag effect of meteorological factors on daily new confirmed cases, the moving average sequences from 1 to 30 days and the values from 1 to 30 days ago of meteorological variables were computed respectively. For example, a family of variables can be derived from temperature average $T$ as follows.

$$T_{\text{avg}}(k) = \frac{1}{s} \sum_{i=1}^{s} T(k-i), 1 \leq s \leq 30 \quad (1)$$

where $T(k)$ represents the temperature average on day $k$ and $s$ represents the forward length of a moving average windows. $T_{\text{avg}}(k)$ and $T_{\text{ago}}(k)$ are the moving average within $s$ days and the value of $s$ days ago of $T(k)$ respectively.

Derived variables of $T_{\text{MAX}}, T_{\text{MIN}}, W$, and $P$ were processed in the same way.

### 2.3. Spearman’s Correlation Coefficient

Spearman’s correlation coefficient is a nonparametric test without prior knowledge, so it is suitable for the correlation test of variables with monotone function relationships. Obviously, the meteorological data used in this study are not normally distributed, and it is more reasonable to assume that there is a monotonic function rather than a linear function relationship between meteorological variables and the spread rate of COVID-19, so the Spearman’s correlation coefficient was used rather than Pearson’s correlation coefficient. Take $T_{\text{avg}}$ and $T_{\text{ago}}$ for example, for the range of $1 \leq s \leq 30$, Spearman’s correlation coefficients were calculated and tested as follows.

$$\rho_{T_{\text{avg}}}(s) = \text{cor}(C(k), T_{\text{avg}}(k)) \quad (3)$$

$$\rho_{T_{\text{ago}}}(s) = \text{cor}(C(k), T_{\text{ago}}(k)) \quad (4)$$

where $C(k)$ denotes daily new confirmed cases on day $k$, and the start date ($k=1$) was defined as the day that total cases exceeded 100 for eliminating the impact of sporadic data. $\rho_{T_{\text{avg}}}(s)$ is the correlation coefficient between sequences of $C(k)$ and $T_{\text{avg}}(k)$, which is a function of window length $s$. $\rho_{T_{\text{ago}}}(s)$ is defined similarly. Other correlation coefficients were calculated in the same way.

### 3. Results and Discussion

According to the climate and land cover type of study area, the 23 counties are divided into 6 zones, as Table 2 shows. Representative counties from 6 zones were picked out to explore the lag effects of meteorological variables on new daily confirmed COVID-19 cases.

Figure 2 shows the trend of Spearman’s correlation coefficients between temperature factors mentioned above and daily confirmed cases for 6 representative counties when the moving average window varies from 1 to 30 days. Figure 2 also shows that correlation coefficients of single day variables (diagonal area) are significantly lower than moving
average variables (grey area), which indicates that moving average variables reflect the changes of daily new cases better. Furthermore, the fluctuations of $T_{\text{avg}}$, $T_{\text{MAX}_{\text{avg}}}$, $T_{\text{MIN}_{\text{avg}}}$ are consistent in a narrow range, and $T_{\text{avg}}$ performs stably. When $s=10$, the correlation between $T_{\text{avg}}$ and new cases tends to be saturated or even to decrease. The $T_{\text{avg}10}$ (moving average of $T$ with a window of 10 days, marked with red triangle) is a good indicator to reflect the daily new cases according to the 6 representative counties in figure 2. In the 6 counties, the correlation in New York County is significantly negative while in other counties are all positive, which indicates that higher temperature can not only restrain COVID-19 but also inspire it in many areas.

Figure 2. Temperature lag effects of Spearman’s correlation coefficient in 6 representative counties with moving-average-windows that varied from 1 to 30 days. The best window lengths for the counties in study area are noted by the red triangle.

Figure 3 shows the lag effects of $W_{\text{avg}}$ and $W_{\text{ago}}$ on COVID-19 over days in the 6 representative counties. The correlation between moving average of wind speed and new cases (red solid line) is stronger than that of single day ago (purple dashed line). As the length of moving average window increase, the correlation performs better. The correlations are significant except in the county of King and Los Angeles. The length of 28 days had the strongest correlation for all concerned, so the $W_{\text{avg28}}$ is chosen to be the indicator.
Figure 3. Wind speed lag effects on Spearman’s correlation coefficient in 6 representative counties with moving-average-windows that varied from 1 to 30 days. The best window lengths for the counties in study area are noted by the red triangles.

Figure 4 shows the lag effects of $P_{\text{avg}}$ and $P_{\text{ago}}$ on COVID-19 over days from the 6 representative counties. The correlation between precipitation and daily new cases is not significant in New York and Salt Lake County. From the other counties we can draw the similar conclusion that the correlation of moving average precipitation (blue solid line) is significant and obviously stronger and more stable than that of single day ago (green dashed line), and the $P_{\text{avg28}}$ is the better indicator.
Based on the analysis from Figure 2 to Figure 4, $T_{avg10}$, $W_{avg28}$ and $P_{avg28}$ were chosen as the best correlation indicators between meteorological variables and the outbreak of COVID-19 in US. The correlation coefficient and significance between the 3 meteorological variables and new daily cases listed in Table 2 shows the selected meteorological indicators perform pretty well in almost all the 23 counties in the study area.

Table 2 shows the details of the 6 climate zones and Spearman’s correlation coefficients between COVID-19 daily new confirmed cases and weather variables. Zone 1 (Climate of temperate deciduous broad-leaved forest) has four distinct seasons, and it is hot and rainy in summer while cold and dry in winter. The temperature indicator in 9 counties from Zone 1 on the east coast is negatively correlated with new cases, which means infection rate of the coronavirus shows a downward trend after the temperature rises in summer. Milwaukee and Hennepin are significantly positively correlated with new cases, because they are located in the interior of US mainland. Wind speed and precipitation are positively correlated with daily new cases in most area, indicates that the persistent wind and precipitation in the past 28 days are good conditions for COVID-19 transmission.
Table 2. Climate Zone and Spearman's correlation coefficients between COVID-19 daily new confirmed cases and 3 weather indicators in 23 counties.

<table>
<thead>
<tr>
<th>Zone</th>
<th>Description</th>
<th>County</th>
<th>( T_{\text{avg}10} )</th>
<th>( W_{\text{avg}28} )</th>
<th>( P_{\text{avg}28} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Zone 1</td>
<td>Climate of temperate deciduous broad-leaved forest</td>
<td>Bergen</td>
<td>-0.704***</td>
<td>0.597***</td>
<td>0.068</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Cook</td>
<td>-0.299***</td>
<td>0.315***</td>
<td>0.235**</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Fairfield</td>
<td>-0.63***</td>
<td>0.689***</td>
<td>0.322***</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Hennepin</td>
<td>0.439***</td>
<td>-0.401***</td>
<td>0.3***</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Marion</td>
<td>-0.346***</td>
<td>0.27***</td>
<td>-0.188*</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Middlesex</td>
<td>-0.614***</td>
<td>0.749***</td>
<td>0.749***</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Milwaukee</td>
<td>0.705***</td>
<td>-0.531***</td>
<td>0.626***</td>
</tr>
<tr>
<td></td>
<td></td>
<td>New York</td>
<td>-0.791***</td>
<td>0.77***</td>
<td>0.047</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Philadelphia</td>
<td>-0.453***</td>
<td>0.442***</td>
<td>0.008</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Providence</td>
<td>-0.402***</td>
<td>NA</td>
<td>0.413***</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Wayne</td>
<td>-0.375***</td>
<td>0.21***</td>
<td>-0.015</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Davidson</td>
<td>0.684***</td>
<td>-0.660***</td>
<td>-0.347***</td>
</tr>
<tr>
<td></td>
<td></td>
<td>District of Columbia</td>
<td>0.137</td>
<td>NA</td>
<td>-0.019</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Fairfax</td>
<td>-0.366***</td>
<td>0.501***</td>
<td>0.098</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Harris</td>
<td>0.782***</td>
<td>-0.681***</td>
<td>0.661***</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Mecklenburg</td>
<td>0.727***</td>
<td>-0.642***</td>
<td>0.009</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Miami-Dade</td>
<td>0.758***</td>
<td>-0.863***</td>
<td>0.294***</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Prince George's</td>
<td>-0.32***</td>
<td>NA</td>
<td>0.133</td>
</tr>
<tr>
<td>Zone 2</td>
<td>Climate of subtropical evergreen broad leaved forest</td>
<td>Clark</td>
<td>0.792***</td>
<td>0.127</td>
<td>-0.464***</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Maricopa</td>
<td>0.778***</td>
<td>0.856***</td>
<td>-0.559***</td>
</tr>
<tr>
<td>Zone 3</td>
<td>Climate of subtropical desert and grassland</td>
<td>Salt Lake</td>
<td>0.752***</td>
<td>-0.191*</td>
<td>0.005</td>
</tr>
<tr>
<td></td>
<td></td>
<td>King</td>
<td>0.20***</td>
<td>0.067</td>
<td>-0.522***</td>
</tr>
<tr>
<td>Zone 4</td>
<td>Temperate desert Climate</td>
<td>Los Angeles</td>
<td>0.746***</td>
<td>0.076</td>
<td>-0.697***</td>
</tr>
</tbody>
</table>

\( \text{avg10} \): average wind speed for March, April, and May; \( \text{avg28} \): average wind speed for June 1–28; \( P \): p-value.

** Table 2. Climate Zone and Spearman's correlation coefficients between COVID-19 daily new confirmed cases and 3 weather indicators in 23 counties.

** Zone 2 (Climate of subtropical evergreen broad leaved forest) contains 7 counties and is very hot and rainy in summer. Fairfax and Prince George’s are contiguous to Zone 1 and have the negative correlation on temperature indicator. The rest of the counties except District of Columbia hold significant positive correlations, which mean the higher temperature in Zone 2 will help COVID-19 to spread. Unlike Zone 1, the wind speed is negatively correlated with new cases in most area. The correlation between outbreak of COVID-19 and meteorological factors in District of Columbia (DC) is not significant, and notice that DC is a famous holiday resort with high population density and high mobility, which may play a more important role in the transmission of COVID-19 than weather factors do. In Zone 3, the temperature and wind speed are significantly positively correlated with daily new cases while precipitation holds a negative correlation, which shows that in dessert and grassland area the higher temperature, higher wind speed and less rainfall will increase the probability of coronavirus infection. In Zone 4, only temperature plays a significant role in the transmissions of coronavirus in Salt Lake County. By contrast, in King County from Zone 5, the correlation between temperature and COVID-19 is not strong as in Salt Lake County, and the precipitation is the strongest indicator in King County. In Zone 6 (Mediterranean climate), it is hot and dry in summer while mild and rainy in winter, and higher temperature in a dry climate will contribute to the outbreak of COVID-19 in the summer in Los Angeles. The temperature holds a negative correlation with coronavirus in Zone 1 and a positive one in other zones. The coronavirus was inactivated more rapidly in higher atmosphere temperature conditions [11]. Additionally, with those conditions, the density of residents is lower in Zone 1, which could lead to the negative correlations. The activities of humans are also an important factor in the spread of coronavirus. For example, higher temperatures may motivate people to take part in activities outdoors, instead of staying at home. This behavior may increase the probability of infection in other Zones especially Zone 2, and the higher population density can amplify the aggregation contact especially in the holiday resort areas such as Miami. As for wind speed, although health officials and logic seem to say that windy conditions would decrease risk of transmission, Table 2 shows that wind speed can increase or decrease the virus transmissions regionally, which depends on whether people get together in close contact outside. In New York County, Our results show that the correlation between COVID-19 and wind speed is positive. The longer the time lag, the more significant the correlation, which is consistent with Muhammad’s and Mehme’s result [8, 22]. The precipitation holds a significant positive correlation with COVID-19 pandemic in more than half of the counties in study area, which means rain helps spreading the coronavirus in most area in US. Conventional wisdom is that COVID is spread indoors much more than outdoors, and rainy days occurred more in early months of the year when people are indoors most of the time. In previous study, the humidity is also associated with
higher risk of COVID-19 and influenza virus [19, 20]. We didn’t consider humidity in this study for lack of the humidity data in the meteorological stations, but precipitation can play a significant role as an indicator for COVID-19 instead of humidity.

Figure 5 shows the spatial distribution of correlation between COVID-19 and meteorological variables. As the green bubbles in Figure 5(a) shows, the temperature indicator $T_{avg10}$ is negatively correlated with COVID-19 in Northeastern United States, while the other area holds a positive correlation (red bubbles). It indicates that hot weather restrains coronavirus infection in Northeastern temperate area and contributes to infection in other area.

Figure 5(b) shows that in Southeastern United States, there is a negative correlation between wind speed indicator $W_{avg28}$ and coronavirus, and other area has a positive one. It means persist wind in the past days is COVID-19’s favorable factor except southeast subtropical area in US.

Figure 5(c) shows precipitation indicator $P_{avg28}$ plays a positive role in COVID-19 pandemic on the east coast of USA and a negative one on the west coast. In the interior of the continent, the effects of precipitation on coronavirus are not clear compared to that on the edge of continent.

Ramadhan obtained a positive correlation between temperature and COVID-19 in Indonesia, while Mehmet got a negative one in Turkey [16, 22]. The results of this study indicate that the influence of meteorological variables on COVID-19 can vary in different countries even regions, because the climate factors and land cover types provide different context.

It was not only the environmental factors, but also the human solutions and habits that made significant influences on the coronavirus transmission. Public health policies of government departments play a key role in the development of the COVID-19 epidemic. For example, New York took measures to close schools, museums, sports events and other similar activities, encouraged people to wear masks to protect themselves and work online as far as possible, and even took curfew measures in a serious situation, and then good effect of coronavirus suppression has been achieved at present.

4. Conclusions

This study analyzes the effects of weather on COVID-19 in temporal and spatial scale in US counties. We examined the associations between meteorological variables and COVID-19 daily new confirmed cases. The findings show that temperature, wind speed and precipitation significantly correlate to the outbreak of COVID-19, negatively or positively, depending on the climate and land cover type.

In the Spearman’s correlation test, the moving average meteorological variables performed better than single day variables. This research shows that the average of temperature within past 10 days, the average of wind speed and precipitation within past 28 days can achieve the optimum.

Temperature was negatively correlated to outbreak of COVID-19 in Northeastern US, and positively in other area. Wind speed was negatively correlated to COVID-19 pandemic in Southeastern US while positively in other area. Precipitation held positive correlation on the east coast of US and negative correlation on the west coast.
continental United States, precipitation held positive or negative correlation with COVID-19, depending on the geographical position.

This study can provide decision makers with valuable information that meteorological factors are significantly associated with the transmissibility of COVID-19, and local health policies should be adopted and implemented according to weather conditions and spatial location.

Acknowledgements

This research was funded by China University of Labor Relations, grant number 20XYJS005. We thank George Hendrey for his valuable comments that improved the manuscript. We also thank the National Oceanic and Atmospheric Administration and the Center for Systems Science and Engineering at Johns Hopkins University for providing the data.

References


[21] Eriko Kudo; Eric Song; Laura J. Yockey; Tasfia Rakib; Patrick W. Wong; Robert J. Homer; and Akiko Iwasaki. Low ambient humidity impairs barrier function and innate resistance against influenza infection. PNAS. 116 (22), 10905–10910. https://doi.org/10.1073/pnas.1902840116.