Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.
The influence of weather conditions on the COVID-19 epidemic
Evidence from 279 prefecture-level panel data in China

Ruofei Lin¹, Xiaoli Wang¹, Junpei Huang*

School of Economics and Management, Tongji University, China

ARTICLE INFO
Keywords:
The COVID-19
Weather conditions
Two-way fixed effect
Heterogeneity analysis
Moderating effects

ABSTRACT:
Studying the influence of weather conditions on the COVID-19 epidemic is an emerging field. However, existing studies in this area tend to utilize time-series data, which have certain limitations and fail to consider individual, social, and economic factors. Therefore, this study aimed to fill this gap. In this paper, we explored the influence of weather conditions on the COVID-19 epidemic using COVID-19-related prefecture-daily panel data collected in mainland China between January 1, 2020, and February 19, 2020. A two-way fixed effect model was applied taking into account factors including public health measures, effective distance to Wuhan, population density, economic development level, health, and medical conditions. We also used a piecewise linear regression to determine the relationship in detail. We found that there is a conditional negative relationship between weather conditions and the epidemic. Each 1 °C rise in mean temperature led to a 0.49% increase in the confirmed cases growth rate when mean temperature was above 7 °C. Similarly, when the relative humidity was greater than 46%, it was negatively correlated with the epidemic, where a 1% increase in relative humidity decreased the rate of confirmed cases by 0.19%. Furthermore, prefecture-level administrative regions, such as Chifeng (included as “warning cities”) have more days of “dangerous weather”, which is favorable for outbreaks. In addition, we found that the impact of mean temperature is greatest in the east, the influence of relative humidity is most pronounced in the central region, and the significance of weather conditions is more important in the coastal region. Finally, we found that rising diurnal temperatures decreased the negative impact of weather conditions on the spread of COVID-19. We also observed that strict public health measures and high social concern can mitigate the adverse effects of cold and dry weather on the spread of the epidemic. To the best of our knowledge, this is the first study which applies the two-way fixed effect model to investigate the influence of weather conditions on the COVID-19 epidemic, takes into account socio-economic factors and draws new conclusions.

1. Introduction

The first case of Corona Virus Disease 2019 (COVID-19; also known as severe acute respiratory syndrome coronavirus 2 or SARS-CoV-2), was identified in December 2019. In just two months, it had spread rapidly and evolved into a global public health crisis. On March 11, 2020, the World Health Organization (WHO) declared the disease a global pandemic (Allam, 2020). As of August 6, 2021, the cumulative number of confirmed COVID-19 cases worldwide exceeded 200 million and the cumulative number of deaths exceeded 4.25 million. The pandemic led to a global crisis that collapsed the world economy to levels not seen since the 1918 influenza pandemic (Zoran et al., 2022). Moreover, more new transmissible variants of SARS-COV-2 have recently been identified (WHO, 2021a, 2021b), such as the Delta variant (B.1.617.2), which has spread globally. The increased incidence and severity of these variants raise grave risks to human health (Bakhshandeh et al., 2021; Grubaugh et al., 2020) and add to the difficulty of implementing pharmacological interventions. There is growing evidence that COVID-19 has the potential to become a seasonal disease like influenza (Lavine et al., 2021), which emphasizes the importance of understanding the mode of transmission of SARS-CoV-2 under environmental conditions such as weather. Exploring the relationship between weather conditions and the spread of COVID-19 is crucial for predicting future outbreaks as well as developing early warning systems, infection control approaches, and public health interventions.

Weather variables are often considered to be key factors in infectious
Some literature suggests that temperature is positively correlated with temperatures resulting in fewer confirmed cases (Nevels et al., 2021; that temperature negatively affects COVID-19 spread, with higher conditions may affect the COVID-19 epidemic through changes in transmission dynamics, host susceptibility, and virus survival in the environment. However, since the COVID-19 virus is very different from known viruses in terms of pathogenicity and transmission, the influence of weather conditions on the COVID-19 epidemic is still an emerging research field, and the existing literature is still relatively limited. Because different research conclusions remain, it is indispensable to carry out further related research. Since the outbreak of the pandemic, extensive research has discussed the influence of weather conditions on the spread of COVID-19 globally, including locations such as China (Shi et al., 2020), the United States (Bashir et al., 2020), Spain (Zoran et al., 2022), and Thailand (Sangkham et al., 2021). Most of them focus on temperature and humidity while a few address other weather conditions such as wind speed and precipitation. Nevertheless, research in this field is in its infancy, there are few relevant papers, and the evidence is very limited (McClymont and Hu, 2021).

Thus, there is still significant debate regarding the impact of temperature and humidity on the COVID-19 pandemic. Most of the existing literature concerning temperature takes minimum, maximum, or mean temperature as weather variables with quite different findings. Much of the literature on this discussion suggests a correlation between temperature and the spread of COVID-19. Although most literature believes that temperature negatively affects COVID-19 spread, with higher temperatures resulting in fewer confirmed cases (Nevels et al., 2021; Hossain et al., 2021; Meyer et al., 2020; Shi et al., 2020; Liu et al., 2020), some literature suggests that temperature is positively correlated with COVID-19 transmission (Sangkham et al., 2021; Iqbal et al., 2020; Islam et al., 2021; Pani et al., 2020). Only a few studies deny any correlation between temperature and outbreaks (Jahangiri et al., 2020; Briz-Redón and Serrano-Aroca, 2020). Regardless of the relationship between humidity and the spread of COVID-19, the existing literature primarily uses absolute or relative humidity as weather variables and shows no agreement on the findings. Most literature reports a negative relationship between humidity and COVID-19 transmission (Wu et al., 2020; Qi et al., 2020; Zhu et al., 2020). However, other studies suggest a positive or no correlation between these factors (Pan et al., 2021; Chien and Chen, 2020).

Some issues are still not well addressed in the existing literature. With regard to sample selection, most of the existing literature selects a single country or region with COVID-19 cases, and the data structure is a time series (Zoran et al., 2022; Islam et al., 2021; Huang et al., 2020; Qi et al., 2020; Ward et al., 2020). However, selecting geographic areas with cases as samples for statistical analysis is prone to sample selection bias (Chen and Astebro, 2001). To accurately verify the relationship between weather conditions and infectious diseases, a study requires the geographic area where the cases occurred and a control group without infectious diseases (Polgreen and Polgreen, 2018). In addition, using only a country-level sample fails to capture regional differences in weather conditions between countries with large areas and uneven population distributions, such as the United States, China, and Brazil. In terms of statistical analysis methods, the existing literature mainly utilizes Spearman’s correlation analysis (Alkhawald et al., 2020; Bashir et al., 2020; Zeng et al., 2020; Joseph et al., 2020). Pearson’s correlation analysis (Islam et al., 2021; Merai et al., 2020), the generalized linear model (GLM; Liu et al., 2020), and the generalized additive model (GAM; Adekunle et al., 2020; Chien and Chen, 2020; Goswami et al., 2020). However, these methods have certain limitations and may lead to bias in estimation due to the inability of the data to satisfy the method’s underlying assumptions.

In addition to the above issues, the most important one not yet addressed in the existing literature is the near absence of consideration of how individual, social, and economic factors influence the spread of COVID-19. Factors such as human behavior patterns (Shammi et al., 2020b; Brockmann and Helbing, 2013), public health measures (Lin et al., 2021), and socio-economic conditions (Jahangiri et al., 2020; Shammi et al., 2020a) are strongly associated with the spread of COVID-19, among which government’s involvement is the most prominent. The COVID-19 epidemic has prompted a broad variety of responses from governments, including, but not limited to, school closures, travel restrictions, prohibitions on public meetings, contact tracing, and contact information sharing (Hale, 2020). These factors have been more influential than weather and climate variables (Oliveiros et al., 2020). Moreover, the intensity of government intervention varies greatly between countries and regions, especially in China, which has taken unprecedentedly comprehensive, strict, and rapid preventive and control measures during this epidemic. However, due to data structure and methodology limitations, most existing studies did not consider these factors as control variables in statistical analysis. The omission of such critical variables will lead to a significant bias in the estimation results (Angrist and Pischke, 2008). The reasons for the spread of COVID-19 are complicated and involve important factors in addition to weather conditions. Therefore, all such factors must be controlled to more effectively analyze the impact of weather on the COVID-19 epidemic.

This study intends to close this research gap. In this article, we collected prefecture-daily panel data for 279 prefecture-level administrative regions in mainland China from January 1 to February 19, 2020, applied two-way fixed effect model of econometrics, and took into account factors such as prevention and control measures taken in China, effective distance to Wuhan (which is the city with the most severe epidemic in China), population density, level of economic development, and health care conditions, to empirically test the effect of weather variables on the COVID-19 confirmed cases growth rate. Unlike other studies, this work is motivated to address the following issues: First, we examine the effects of mean temperature and relative humidity on the confirmed cases growth rate. Then we define “dangerous weather” based on the previous results, and identify “warning cities” that are most likely to have outbreaks based on the number of days with dangerous weather. Then, we explore the regional heterogeneity of weather effects. Finally, we discuss the moderating effects of diurnal temperature differences, public health measures, and social opinion on the effects of weather conditions by introducing interaction terms to examine the influence of these factors on the impact strength of weather conditions.

2. Materials and methods

2.1. Study area

Our study included 279 prefecture-level administrative regions in China, covering most of mainland China. Fig. 1 illustrates the number of confirmed cases in China as of February 19, 2020.

China’s territory spans tropical, temperate, and frigid zones from south to north, where the Qinling-Huaihe line, as a geographic boundary, is 0 °C isotherm in January. That same month the mean temperature is above 0 °C south of the Qinling-Huaihe line and below 0 °C in the north. There is a big difference in temperature between the northern and southern prefecture-level administrative regions. In January 2020 the average high temperature in the northern city of Heihe was −13 °C and the average low temperature was −25 °C, while in the southern city of Sanya, it was 25 °C and 16 °C with a difference of 38 °C and 41 °C, respectively. In February 2020 the average high temperature in Heihe was −9 °C and the average low temperature was −21 °C, while in Sanya, it was 27 °C and 19 °C with a difference of 36 °C and 40 °C, respectively. In China, January is usually the coldest month, but the temperature starts to increase in February and rises sharply in March when spring begins. At that point the country’s temperatures are mostly above 0 °C and temperature differences between regions are greatly reduced. In
March 2020 the highest temperature in Sanya reached 30 °C; that in Heihe was 14 °C. The regional temperature difference between the north and south of China provides an excellent case for studying the influence of weather conditions on the COVID-19 epidemic, especially the large temperature differences in January and February between cold temperatures below 0 °C in the north and warm temperatures above 0 °C in the south.

In addition, China also adopted stringent public health measures that brought the epidemic under control within a little more than two months. On February 12 the number of new cases reached a peak of 15,152. On February 19 the number of nationwide cases dropped to 394; except for a few provinces, such as Hubei, most others had very low case numbers. For the first time on March 20, there were no new confirmed cases, and after that date most provinces had no new cases or only sporadic cases (see Fig. 2). In this paper, considering the temperature and the number of confirmed cases simultaneously, we selected the data set of all prefecture-level administrative regions in mainland China from January 1 to February 19, 2020, as the research sample.

2.2. Data collection

This paper selected a dataset of 279 prefecture-level administrative regions in mainland China from January 1 to February 19, 2020, as the study sample. Daily confirmed cases were collected from the official release of the national and local Health Commission. Weather variables were collected from the website of the National Meteorological Information Center (NMIC). Weather factors included daily mean temperature and relative humidity.

In terms of control variables, we first constructed the scoring data representing the intensity of prevention and control by referring to Lin et al. (2021) method as a proxy variable of public health measures. The comprehensive intervention measures implemented in China successfully alleviated the spread of COVID-19, especially in the early stages of the outbreak, so we included that factor in the model. We conducted a very detailed evaluation through the manual collection of information or announcements issued by the prevention and control headquarters of the prefecture-level administrative districts (see Data Supplement A1).

Brockmann and Helbing (2013) proposed a new concept: effective distance. They believed that the spread of disease has nothing to do with the geographic distance between cities but is closely related to the effective distance. Therefore, we calculated the effective distance from each prefectural-city to Wuhan, the city with the worst outbreak in China, by referring to Brockmann and Helbing (2013). The data of population movement used to calculate the effective distance comes
from Baidu Migration, which is a website release the data of migration and outflow between different regions collected by Baidu Inc using big data technology (see Data Supplement A2).

Other control variables included number of beds in medical institutions, GDP per capita, and population density collected from the China City Statistical Yearbook. Due to the lack of statistical data in some prefecture-level administrative regions in this resource, the final number of regions used was 279. Variable explanation and data source were shown in Table 1.

2.3. Calculation of actual cumulative case growth rate

The explained variable in this paper is actual cumulative case growth rate; it can be described as follows:

\[
\text{Actual cumulative case growth rate}_i(t) = \left( \frac{\text{Actual cases}_i - \text{Actual cases}_{i, t-1}}{\text{Actual cases}_{i, t-1}} \right) 
\]

If there are no cases on the current day, \( \text{Actual cases}_{i, t-1} = 0 \); therefore, let \( \text{Actual cumulative case growth rate}_i(0) = 0 \).

If there are no confirmed cases on that day, \( \text{Actual cumulative cases}_{i, t-1} = 0 \); therefore, let the cumulative case growth rate be 0.

\[
\text{rate}_i = \alpha + \beta_1 f(\text{Weather}_i) + \beta_2 \text{Measure}_i + \beta_3 \ln \text{bed}_i + \beta_4 \ln \text{pop}_i + \beta_5 \ln \text{gdp}_i + \beta_6 \text{distance}_i + \tau + \delta_i + \tilde{\delta}_i + \epsilon_i 
\]

Considering that the average incubation period of COVID-19 is 5.2 days (H. Li et al., 2020), we take the fifth lead of reported cases as the proxy variable of the actual cases, namely:

\[
\text{reported case}_i = \text{actual case}_i,_{t-5} 
\]

2.4. Statistical analysis

We applied the econometrics approach to empirically test the influence of weather conditions on the epidemic. The econometric approach is commonly used to measure the effects of a factor on economic growth. Similar to early COVID-19 infections, economic output generally increases exponentially with a variable rate that can be affected by policies and other conditions (Hsiang et al., 2020). Therefore, it is appropriate to apply econometrics techniques to analyze the influence of weather condition on the outbreak of the epidemic. Compared to statistical methods such as Pearson’s correlation coefficient to identify the correlation between weather and COVID-19 (Alkhouwail et al., 2020; Bashir et al., 2020; Menebo, 2020), using the multiple linear regression approach of panel data is more conducive to overcoming the limitations of the time-series regression and the Spearman regression model adopted by most existing literature. The econometric approach pays more attention to identifying the causal relationship between the variables (Angrist and Pischke, 2008), that is, whether changes in temperature and humidity lead to changes in the epidemic spread. For causal inference the key is to control the observable factors that interfere with the causal relationship. In order to avoid biased estimators created by omitting variables, we adopt the two-way fixed effect model to control the time-invariant individual heterogeneity and the individual-invariant time heterogeneity. The empirical model is as follows:

\[
\text{rate}_i = \alpha + \beta_1 f(\text{Weather}_i) + \beta_2 \text{Measure}_i + \beta_3 \ln \text{bed}_i + \beta_4 \ln \text{pop}_i + \beta_5 \ln \text{gdp}_i + \beta_6 \text{distance}_i + \tau + \delta_i + \tilde{\delta}_i + \epsilon_i 
\]

where \( \text{rate}_i \) is the explained variable, representing the actual cumulative case growth rate of city \( i \) on date \( t \).

\( f(\text{Weather}_i) \) denotes the polynomial of weather variables, including \( \text{MT}_{c_i} \) denoting mean temperature of city \( i \) on date \( t \), and \( \text{RH}_{c_i} \) denoting relative humidity of city \( i \) on date \( t \).

\( \text{Measure}_i \) denotes the total score of public health measures of city \( i \) on date \( t \), \( \text{distance}_i \) denotes the effective distance from city \( i \) to Wuhan on date \( t \), and \( \ln \text{bed}_i \), \( \ln \text{pop}_i \), \( \ln \text{gdp}_i \) denotes number of beds in medical institutions, GDP per capita, and population density, respectively. \( \delta_i \) is a region fixed effect to control the characteristics of provinces constant over time, \( \tilde{\delta}_i \) is a time fixed effect to control the time factors that do not vary from individual to individual, \( \epsilon_i \) is an error term, and we use cluster-robust standard error to estimate the standard deviation (Cameron and Miller, 2015).

According to the results from the baseline regression, we used a piecewise linear regression to determine in detail the relationship between weather variables and COVID-19 confirmed cases’ growth rate (Xin and Zhu, 2020; Kim et al., 2016). In the sensitivity analysis, we first tried to exclude Hubei province, the most severely affected province in China, from the full sample. We then adjusted the assumptions of the length of the incubation period when calculating actual confirmed cases’ growth rate to examine whether our main results were robust. Based on the statistical analysis results above, we defined “dangerous weather” (that which is most conducive to the development of the epidemic), counted the frequency of dangerous weather in various Chinese cities, and identified the “warning cities” most likely to experience outbreaks.

Further, we explored the regional heterogeneity of the effect of weather on the COVID-19 epidemic. We conducted subsample regressions according to the characteristics of China’s geographical conditions in the east-central-west and coastal-inland areas, respectively. Finally, we discussed the moderating effects of diurnal temperature
differences, public health measures and social opinion on the impact of weather conditions by introducing interaction terms to examine the influence of these three factors on the impact strength of weather conditions.

We used STATA (version 15.0) in this study to conduct the analysis. The statistical tests were two-sided, and the p-value told us if the result were statistically significant. P-values of $<0.01$ (**), $<0.05$ (*) and $<0.1$ (*) were considered statistically significant.

3. Result and discussion

3.1. Descriptive analysis

The data samples in this paper are composed of balanced panel data of 279 prefecture-level administrative regions from January 1 to February 19, 2020, and the descriptive statistics of related variables are shown in Table 2.

Fig. 3 shows the distribution of the mean value of the mean temperature from January 1 to February 19, 2020, from which it can be seen that the mean temperature decreases in a step-like manner from south to north. Fig. 4 shows the distribution of the mean of the relative humidity from January 1 to February 19, 2020. It can be seen that the relative humidity in the southeast coastal area is higher, and it shows a decreasing trend in the northwest direction.

3.2. Influence of weather conditions on confirmed cases growth rate: conditional negative linear relationship

Table 3 reports the results of the baseline regression. The explanatory variable is the growth rate of cumulative cases, and the explanatory variable is the polynomial of mean temperature and relative humidity. Both the first and quadratic terms for mean temperature and relative humidity are significantly negative, which suggests a nonlinear relationship between mean temperature, relative humidity, and cumulative case growth rates. Specifically, the relationship was approximately positive linear when the mean temperature was $<7 \degree C$ and became negative linear above $7 \degree C$, indicating that the single threshold of the temperature effect on COVID-19 was $7 \degree C$. Similarly, the threshold for relative humidity was 46%.

We also analyze control variables. It is observed that the coefficient of public health measures is significantly negative, indicating that taking public health measures are essential to mitigating the epidemic. The better the public health measures are, the lower the number of confirmed cases growth rate is. The coefficient of population size was significantly positive. The vast population size will increase the difficulty of isolating person-to-person contact, which adversely effects blocking the further spread of infectious diseases. The coefficients of both the number of beds in health institutions and effective distance were significantly negative. The number of beds in health institutions represents the condition of the city’s medical resources, and cities with

| Variables | Implied | Notation | N  | mean  | sd   | min  | max  |
|-----------|---------|----------|----|-------|------|------|------|
| Explained variable | Increased rate of confirmed cases | Rate | 13,950 | 0.0851 | 0.443 | 0 | 19 |
| Explanatory variable | mean temperature | MT | 13,950 | 3.789 | 9.0667 | -31.2 | 26.4 |
| | Relative humidity | RH | 13,850 | 71.17 | 17.15 | 6 | 102 |
| Control variable | Public health measures score | Measure | 13,950 | 3.729 | 3.912 | 0 | 10 |
| | Number of registered population | Pop | 13,950 | 171.3 | 226.3 | 16 | 2451 |
| | Number of hospital beds | Hospital_bed | 13,950 | 12,906 | 17,135 | 920 | 142,708 |
| | GDP per capita | Per GDP | 13,950 | 92,348 | 379,890 | 17,890 | 6.400e+06 |
| | effective distance | Distance | 13,950 | 5.722 | 1.874 | 0 | 7.785 |

Table 2: Statistical description.

Fig. 3. Mean Value of the mean temperature in China from January 1 to February 19, 2020.
richer medical resources are more capable of mitigating the impact of
the epidemic. The shorter the effective distance to Wuhan, the more
severe the epidemic’s outbreak, which is in line with theoretical
expectations.

Based on baseline results, a piecewise linear regression was adapted
with a threshold at $7^\circ C$ for mean temperature and 46% for relative
humidity to quantify the effect of weather conditions above and below
the threshold. Table 4 shows the results of the piecewise linear regres-
sion. It can be seen that each $1^\circ C$ rise in mean temperature led to a
0.49% increase in confirmed cases’ growth rate when mean temperature
was above $7^\circ C$, while when mean temperature was below $-7^\circ C$, the
positive effect of temperature was not statistically significant. Similarly,
when the relative humidity was greater than 46%, it was negatively
correlated with the epidemic, with a 1% increase in relative humidity
decreasing the rate of confirmed cases by 0.19%. However, when the
relative humidity was less than 46%, the decrease in humidity no longer
affected the outbreak’s spread. Therefore, we conclude that there is a
conditional linear relationship between the weather variable and
COVID-19 as shown in Fig. 5.

Our results are in good agreement with other studies, concluding that
mean temperature and relative humidity parameters play an important
role in the COVID-19 epidemic (Liu et al., 2020; Liu et al., 2020; Xie and
Zhu, 2020; Tobías and Molina, 2020; Tantrakarnapa et al., 2020; Pani
et al., 2020). Regarding weather temperature, although most findings
suggested a negative correlation between temperature and the
COVID-19 epidemic (Zoran et al., 2022; Nevels et al., 2021; Hossain
et al., 2021; Pahuja et al., 2021; Meyer et al., 2020; Shi et al., 2020; Liu
et al., 2020; Tobías and Molina, 2020; Menebo, 2020; Kumar & Kumar,
2020; Li et al., 2020; Menebo, 2020, 2020; Prata et al., 2020), some
research indicated that temperature positively affects epidemics (Xie
and Zhu, 2020; Iqbal et al., 2020; Islam et al., 2021; Pani et al., 2020;
Tosepu et al., 2020), and a few found no relationship (Jahangiri et al., 2020; Briz-Redón and Serrano-Aroca, 2020).

Also, our findings are consistent with the belief that higher relative humidity can alleviate the COVID-19 outbreak, which was also demonstrated by most of the existing studies examining the same hypothesis with different data and approaches (Wu et al., 2020; Qi et al., 2020; Zhu et al., 2020; Sahin, 2020; Ahlawat et al., 2020). Conversely, other studies revealed a positive correlation (Alkhawalled et al., 2020; Chien and Chen, 2020; Pani et al., 2020), and a few reported no relationship (Pan et al., 2021).

### 3.3. Sensitivity analysis

First, considering that the epidemic first broke out in the city of Wuhan and had a severe impact on other cities in Hubei Province, we eliminated Hubei Province sample. Second, we adjusted the incubation period length hypothesis, assuming that the incubation period is 6 days and 7 days, respectively, and then performed the regression again. The results of the robustness were reported in Table 5. It can be seen that the sign and significance of the coefficients are consistent with the baseline regression, indicating that the conclusion was still robust after changing the sample selection and basic assumptions.

### 3.4. Defining “dangerous weather” and identifying “warning cities”

In China, a county with vast territory and wide differences in climatic conditions, it is necessary to further explore which cities have weather conditions that are more likely to have outbreaks. We defined a temperature of $7^\circ C \pm$ one standard deviation (9.0667) as “dangerous temperature” and a relative humidity of 46% $\pm$ one standard deviation (17.15) as “dangerous humidity.” We counted the number of days with “dangerous weather” in each city in China from January 1 to February 19, 2020. Then, we defined the top 10 cities with the number of days with “dangerous weather” that would designate them as “warning cities,” that is, cities more at risk of COVID-19 outbreaks. Table 6, panels A, B, and C, lists the “warning cities” in China that are the top 10 cities with...
dangerous temperatures, dangerous humidity (level 2), and with both dangerous temperatures and dangerous humidity (level 1), respectively. The results show that, in winter, cities such as Chifeng and Zhangjiakou were shortlisted as early warning cities with weather conditions more favorable for an epidemic outbreak.

Although there have been several studies that explored the impact of weather conditions on the COVID-19 epidemic in China (Shi et al., 2020; Liu et al., 2020; Xie and Zhu, 2020; Ma et al., 2020), these studies only analyzed and discussed the relationship between weather conditions and outbreaks without further indicating which cities in China were more likely to have outbreaks based on the results. Our study filled in the void.

Table 7
Sub-sample results.

| Panel A East Middle and West | (1) | (2) | (3) | (4) | (5) | (6) |
|-----------------------------|-----|-----|-----|-----|-----|-----|
| MT ≥ -7 °C                  |     |     |     |     |     |     |
| East                        | -0.0059*** | -0.0025 | -0.0056*** |     |     |     |
| Middle                      |     |     |     |     |     |     |
| West                        |     |     |     |     |     |     |
| RH                          |     |     |     |     |     |     |
| East                        | -0.0021*** | -0.0025*** | -0.0015*** |     |     |     |
| Middle                      |     |     |     |     |     |     |
| West                        |     |     |     |     |     |     |
| Observations                | 4128 | 3604 | 3257 | 4082 | 4255 | 3040 |
| R-squared                   | 0.054 | 0.041 | 0.040 | 0.054 | 0.041 | 0.047 |
| Control Variables           | YES | YES | YES | YES | YES | YES |
| Time Trend                  | YES | YES | YES | YES | YES | YES |
| Province FE                 | YES | YES | YES | YES | YES | YES |
| Time FE                     | YES | YES | YES | YES | YES | YES |

Panel B Coastal and Inland

| (1) | (2) | (3) | (4) |
|-----|-----|-----|-----|
| MT ≥ -7 °C |     |     |     |
| East | -0.0098** | -0.0044*** |     |
| Middle |     |     |     |
| West |     |     |     |
| RH |     |     |     |
| East | -0.0018* | -0.0017*** | -0.0021* | -0.0017*** |
| Middle |     |     |     |
| West |     |     |     |
| Observations | 2191 | 8798 | 2101 | 9276 |
| R-squared | 0.051 | 0.043 | 0.051 | 0.044 |
| Control Variables | YES | YES | YES | YES |
| Time Trend | YES | YES | YES | YES |
| Province FE | YES | YES | YES | YES |
| Time FE | YES | YES | YES | YES |

Table 8
Exploration of moderating effects.

| (1) | (2) | (3) | (4) | (5) | (6) |
|-----|-----|-----|-----|-----|-----|
| MT |     |     |     |     |     |
| rate | -0.0033*** | -0.0063*** | -0.0056*** |     |     |
| (0.0015) | (0.0014) | (0.0014) |     |     |
| RH |     |     |     |     |     |
| rate | -0.0018*** | -0.0034*** | -0.0022*** |     |     |
| (0.0004) | (0.0004) | (0.0004) |     |     |
| MT*highf |     |     |     |     |     |
| -0.0034*** | (0.0011) | 0.0040*** |     |     |
| (0.0011) | (0.0005) | (0.0013) |     |     |
| RH*highf |     |     |     |     |     |
| -0.0005*** | (0.0001) | 0.0028*** |     |     |
| (0.0001) | (0.0002) |     |     |     |
| MT*highpolicy |     |     |     |     |     |
| 0.0004*** | (0.0001) |     |     |     |
| (0.0001) |     |     |     |     |
| RH*highpolicy |     |     |     |     |     |
| 0.0030*** | (0.0001) |     |     |     |
| (0.0001) |     |     |     |     |
| Observations | 10,989 | 11,377 | 10,989 | 11,377 | 10,989 | 11,377 |
| R-squared | 0.038 | 0.037 | 0.038 | 0.050 | 0.037 | 0.036 |
| Control Variables | YES | YES | YES | YES | YES | YES |
| Time Trend | YES | YES | YES | YES | YES | YES |
| Province FE | YES | YES | YES | YES | YES | YES |
| Time FE | YES | YES | YES | YES | YES | YES |

Standard errors in parentheses, ***p < 0.01, **p < 0.05, *p < 0.1.

Note: According to the classification of the National Bureau of Statistics of China, http://www.stats.gov.cn/, the eastern region includes Beijing, Tianjin, Hebei, Liaoning, Shanghai, Jiangsu, Zhejiang, Fujian, Shandong, Guangdong, and Hainan. The middle region includes Shanxi, Inner Mongolia, Jilin, Heilongjiang, Anhui, Jiangxi, Henan, Hubei, Hunan and Guanxi. The western region includes Sichuan, Guizhou, Yunnan, Tibet, Shanxi, Gansu, Qinghai, Ningxia, and Xinjiang. According to the China Marine Statistical Yearbook, coastal areas are defined as areas with coastlines, which are divided into coastal provinces, autonomous regions and municipalities. At present, there are 53 coastal cities and 242 coastal counties.
3.5. Regional heterogeneity of weather effect

Furthermore, this paper conducted sub-sample regression according to geographical location to explore the heterogeneity of the influence of weather conditions. The results were shown in Table 7. Panel A in Table 7 reports the results according to the samples in the east, middle, and west areas. It can be seen that the coefficient of mean temperature is still significantly negative in the eastern and western regions, where the influence of mean temperature in the east is greater than in the west. However, this doesn’t work in the middle areas. The effect of relative humidity is most significant in the middle, followed by the east, and is the weakest in the west. Panel B in Table 7 reports the sub-sample results of the coastal and inland areas. It shows that the influence of both mean temperature and relative humidity is greater in the coastal areas, and the impact of weather conditions is more significant in coastal areas than inland areas.

Most of the literature exploring the impact of weather conditions on the epidemic in China used time series data (Qi et al., 2020; Shi et al., 2020; Ma et al., 2020) with single samples, such as the whole country or a specific city. Limitations of sample and data structure prevented these studies from further exploring the regional heterogeneity of the weather effect. While there are some studies with samples of multiple cities in China (Liu et al., 2020; Xie and Zhu, 2020) they also don’t discuss regional heterogeneity. In truth, weather conditions vary greatly among regions in China, which are affected not only by large north-south latitude span, but also by significant differences in altitude and the location of coastal and inland areas. Therefore, it is necessary to conduct a heterogeneity analysis on the impact of weather conditions to explore the differences in the impact of mean temperature and relative humidity on the epidemic in various regions. Our study complements this issue.

3.6. Moderating effects of temperature variation, public health measures and social opinion on the impact of weather conditions

We introduced the interaction term of the explanatory variables with diurnal temperature variations, public health measures, and social concern respectively to explore the moderating effects of these factors on weather conditions affecting the epidemic. Data collection for diurnal temperatures difference and social opinion variables are shown in Data Supplement A3.

We divided the three variables of diurnal temperature differences, public health measures, and social concern into high, medium, and low, respectively. We then generated dummy variables, and constructed the interaction terms between the dummy and weather condition variables. The results are shown in Table 8. Column (1) introduces the interaction between mean temperature and high diurnal temperature variations (highT). Column (2) introduces the interaction between relative humidity and the item. Similarly, column (3) and column (4) presents the results of public health measures, and column (5) and column (6) presents the results of social opinion.

It can be seen that, whether involving mean temperature or relative humidity, the coefficient of the interaction with high diurnal temperature differences (highT) is significantly negative, that is, the increase in diurnal temperature differences leads to a stronger impact of weather conditions on the increase in the rate of confirmed COVID-19 cases, especially in dry and cold regions, where higher diurnal temperature differences increase the risk of epidemic spread. On the contrary, the coefficients of the interaction with high public health measures (highpolicy) and high social concern (highconcern) are both significantly positive, indicating that the improvement of public health measures and social concern can weaken the influence of mean temperature and relative humidity on the growth rate of COVID-19 confirmed cases. It can be concluded that strict public health measures and high social concern can mitigate the adverse effects of cold and dry weather on the spread of the epidemic, which reinforces the importance of public health measures and attention to public response.

The main factors influencing temperature differences between day and night are latitude, altitude, and location of the land and sea. A large temperature difference between day and night weakens the immune system and makes people more susceptible to infection under equal conditions (Jaagus et al., 2014). Although the role of temperature differences has been analyzed in the literature (Liu et al., 2020; Park et al., 2020), no studies have focused on the interaction of temperature differences with other weather conditions. In fact, temperature differences should be taken into account even under the same temperature and humidity conditions, especially in areas with dry and cold climates, where higher diurnal temperature differences will exacerbate the risk of outbreak spread.

Studies of the impact of weather conditions cannot occur separately from the effect of public health interventions and human behavior patterns on the spread of the epidemic (Polgreen and Polgreen, 2018; Paraskevis et al., 2020). It has been proved that rapid and strict public health measures can effectively prevent the further spread of the epidemic (Lin et al., 2021; Thu et al., 2020), and high social concern can enhance the public’s attention on the epidemic to improve prevention awareness (Hussain, 2020). However, to date, no studies have further explored the moderating effects of public health measures and social opinion on weather impacts. By exploring the effects of climatic conditions on the COVID-19 pandemic, this study found that the adoption of strict public health measures and the availability of good social opinion can mitigate the adverse effects of dry and cold weather on the spread of the epidemic. These findings provide a realistic argument for stricter preventive and control measures and better social opinion in warning cities with frequent dangerous weather.

4. Conclusions

To the best of our knowledge, this is the first study which applies the two-way fixed effect model using prefecture-daily panel data to explore the impact of weather conditions on the COVID-19 epidemic. It considers important influencing factors such as human behavior patterns, public health measures, and socio-economic conditions, and then draws new conclusions.

Future research can adopt more detailed investigation methods. Under the legal framework of privacy protection, questionnaire surveys could be carried out with patients’ consent to draw more accurate conclusions. At the same time, in terms of the mechanism of the role of weather variables, more in-depth interdisciplinary cooperation with epidemiologists is needed to study the specific impact of weather conditions on the survivability of the COVID-19 virus and the immunity of susceptible populations to obtain a clearer picture and offer compelling conclusions.

To conclude, combating the COVID-19 pandemic is a difficult undertaking. Instead of depending only on changes in weather conditions, a portfolio of tactics should be created. Given the current climate-change framework, it is important to raise the alarm about the possibility of a more severe pandemic in the near future.

Author contributions

Ruofei Lin: Conceptualization, Data curation, Visualization, Formal analysis. Xiaoli Wang: Writing- Reviewing and Editing; Validation, Supervision. Junpei Huang: Methodology, Software, Writing - Original Draft.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.
Acknowledgements

The authors gratefully acknowledge the financial support of the National Social Science Foundation of China (Grant No. 13BJY091) and the National Natural Science Foundation of China (Grant No.71773083). Our deepest gratitude goes to the editor and anonymous reviewers for their careful work and thoughtful suggestions that have helped improve this paper substantially.

References

Adekunle, I.A., Tella, S.A., Oyesiku, K.O., Oseni, I.O., 2020. Spatio-temporal analysis of meteorological factors in abating the spread of COVID-19 in Africa. Heliyon 6 (8), e04749.
Ahlawat, A., Wiedensohler, A., Mishra, S.K., 2020. Impact of meteorological parameters on COVID-19 pandemic: a comprehensive study from Saudi Arabia. Infec. Med. Unlocked 2020, 100418.
Allam, Z., 2020. The first 50 days of COVID-19: a detailed chronological timeline and analysis. Infect. Genet. Evol. 47, 103927.
Angrist, J.D., Pischke, J.S., 2008. Mostly Harmless Econometrics: an Empiricist’s Companion. Princeton university press.
Angrist, J.D., Pischke, J.S., 2008. Mostly Harmless Econometrics: an Empiricist’s Companion. Princeton university press.
Ahlawat, A., Wiedensohler, A., Mishra, S.K., 2020. An overview on the role of relative humidity, temperature and meteorological parameters in COVID-19 pandemic in Mumbai, India. Diabetes Metab. Syndr. 14 (6), 1735–1742.
Bakhshandeh, B., Sorboni, S.G., Javanmard, A.-R., et al., 2021. Variants in ACE2; impact of renin-angiotensin system on COVID-19 transmission in different countries. Front. Cell Infect. Dis. 66 (9), 817.
Brockmann, D., Helbing, D., 2013. The hidden geometry of complex, network-driven epidemics. PLoS ONE 8 (7), e67586.
Brockmann, D., Helbing, D., 2013. The hidden geometry of complex, network-driven contagion phenomena. Sci. Rep. (6164), 1337–1342.
Briz-Redon, A., Serrano-Aroca, Á., 2020. A spatio-temporal analysis for exploring the effect of temperature on COVID-19 early evolution in Spain. Sci. Total Environ. 728, 138811.
Brockmann, D., Helbing, D., 2013. The hidden geometry of complex, network-driven contagion phenomena. Sci. Rep. (6164), 1337–1342.
Baidu Migration website: http://qixian.baidu.com/, access date August 2021.
Bakhshandeh, B., Sorboni, S.G., Javanmard, A.-R., et al., 2021. Variants in ACE2; impact of renin-angiotensin system on COVID-19 transmission in different countries. Front. Cell Infect. Dis. 66 (9), 817.
Brockmann, D., Helbing, D., 2013. The hidden geometry of complex, network-driven contagion phenomena. Sci. Rep. (6164), 1337–1342.
Baker, H.G., 2010. Theor. Appl. Climatol. 118 (1), 57–68.
Baker, H.G., 2010. Theor. Appl. Climatol. 118 (1), 57–68.
Baker, H.G., 2010. Theor. Appl. Climatol. 118 (1), 57–68.
Baker, H.G., 2010. Theor. Appl. Climatol. 118 (1), 57–68.
Baker, H.G., 2010. Theor. Appl. Climatol. 118 (1), 57–68.
Baker, H.G., 2010. Theor. Appl. Climatol. 118 (1), 57–68.
Baker, H.G., 2010. Theor. Appl. Climatol. 118 (1), 57–68.
Baker, H.G., 2010. Theor. Appl. Climatol. 118 (1), 57–68.
Baker, H.G., 2010. Theor. Appl. Climatol. 118 (1), 57–68.
Baker, H.G., 2010. Theor. Appl. Climatol. 118 (1), 57–68.
Baker, H.G., 2010. Theor. Appl. Climatol. 118 (1), 57–68.
Baker, H.G., 2010. Theor. Appl. Climatol. 118 (1), 57–68.
Baker, H.G., 2010. Theor. Appl. Climatol. 118 (1), 57–68.
Baker, H.G., 2010. Theor. Appl. Climatol. 118 (1), 57–68.
Baker, H.G., 2010. Theor. Appl. Climatol. 118 (1), 57–68.
Shammi, M., Bodrud-Doza, Md, Islam, A.R. M. d T., Rahman, M. d M., 2020a. Psychosocial, and socio-economic crisis in Bangladesh due to COVID-19 pandemic: a perception-based assessment. Heliyon 6 (5), e04063.

Shi, P., Dong, Y., Yan, H., Zhao, C., Li, X., Liu, W., et al., 2020. Impact of temperature on the dynamics of the COVID-19 outbreak in China. Sci. Total Environ. 728, 138890.

Tan, J., Ma, L., Huang, J., Yu, S., Chen, B., Yin, J., 2005. An initial investigation of the association between the SARS outbreak and weather: with the view of the environmental temperature and its variation. J. Epidemiol. Community Health 59 (3), 186–192.

Tantrakarnapa, K., Bhopdhornangkul, B., Nakhaapakorn, K., 2020. Influencing factors of COVID-19 spreading: a case study of Thailand. J. Public Health 1–7.

Thu, T.P.B., Ngoc, P.N.H., Hai, N.M., 2020. Effect of the social distancing measures on the spread of COVID-19 in 10 highly infected countries. Sci. Total Environ. 742, 140430.

Tobías, A., Molina, T., 2020. Is temperature reducing the transmission of COVID-19? Environ. Res. 186, 109553.

Tosepu, R., Gunawan, J., Effendy, D.S., Lestari, H., Bahar, H., Aulfian, P., 2020. Correlation between weather and covid-19 pandemic in jakarta, Indonesia. Sci. Total Environ. 725, 138436.

Ward, M.P., Xiao, S., Zhang, Z., 2020. The role of climate during the COVID-19 epidemic in New South Wales, Australia. Transbound Emerg. Dis. 67, 2313–2317.

WHO, 2021a. COVID-19 Weekly Epidemiological Update 12 January 2021 [WWW Document]. 1.13.21. https://www.who.int/docs/default-source/coronaviruse/situation-reports/weekly-epidemiological-update_22.pdf.

WHO, 2021b. https://www.who.int/emergencies/diseases/novel-coronavirus-2019/situation-reports.

Wu, Y., Jing, W., Liu, J., Ma, Q., Yuan, J., Wang, Y., Du, M., Liu, M., 2020. Effects of temperature and humidity on the daily new cases and new deaths of COVID-19 in 166 countries. Sci. Total Environ. 729, 139051.

Xie, J., Zhu, Y., 2020. Association between ambient temperature and COVID-19 infection in 122 cities from China. Sci. Total Environ. 724, 148201.

Zhu, L., Liu, X., Huang, H., et al., 2020. Meteorological impact on the COVID-19 pandemic: a study across eight severely affected regions in South America. Sci. Total Environ. 744, 140881.

Zoran, M.A., Savastru, R.S., Savastru, D.M., Tautan, M.N., Baschir, L.A., Tenciu, D.V., 2022. Assessing the impact of air pollution and climate seasonality on COVID-19 multiwaves in Madrid, Spain. Environ. Res. 203, 111849.