Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

# Exploration of Temporal-Spatially Varying Impacts on COVID-19 Cumulative Case in Texas Using Geographically Weighted Regression (GWR)

**XIU WU**
Texas State University
**Jinting Zhang** ( whuzjt@whu.edu.cn )
Wuhan University https://orcid.org/0000-0003-3037-7645

**Research Article**

**Keywords:** Geographical Weighted Regression (GWR), Temporal-spatially varying impacts, COVID-19 Cumulative Case

**Posted Date:** March 4th, 2021

**DOI:** https://doi.org/10.21203/rs.3.rs-264154/v1

**License:** This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

**Version of Record:** A version of this preprint was published at Environmental Science and Pollution Research on April 10th, 2021. See the published version at https://doi.org/10.1007/s11356-021-13653-8.

# Exploration of Temporal-spatially varying Impacts on COVID-19 Cumulative Case in Texas using geographically weighted regression (GWR)

**Xiu Wu [1], Jinting Zhang [2,*]**

1 Department of Geography, Texas State University; x_w10@txstate.edu; zhan@txstate.edu; chow@txstate.edu

**2*** Correspondence: whuzjt@whu.edu.cn.edu; School of Resource and Environmental Science, Wuhan University; whuzjt@whu.edu.cn Tel.: +86-1338-750-9245 (F.L.)

**Abstract:** Since COVID-19 is extremely menacing human's health, it is a significant to expose on its fator's impacts for curbing the virus spreading. To tackle the complexity of COVID-19 expansion in spatial-temporal scale, This research is approriatedly analyzed the spatial-temporal heterogeneity at county-level in Texas. First,factors impacts of COVID-19 are captured on social, economic, and environmental multiple-facets and the Communality is extracted through Principal Component Analysis (PCA). Second, this research is used COVID-19 CC as the dependent variable and the common factors as the independent variable. According to the virus prevailing hierarchy, spatial-temporal disparity is are categorized four quarters in the modeling GWR analysis according to the virus prevailing hierarchy. The findings are exibited that GWR models provided higher fitness, more geodata-oriented information than OLS models. In Texas El Paso, Odessa, Midland, Randall and Potter County areas, population, hospitalization, and age structure presented static, positive influences on COVID-19 cumulative casesm, indicating they should be adopt stringent stratgies in curbing COVID-19. Winter is the most sensitive season for the virus spreading, implying the last quarter should be pay more attention to prevent the virus and take pracutions. This research are expected to provide references for preventing and controlling COVID-19 and related infectious dieseaces, evidences for disease surveillance and response systems to facilitate the appropriate uptake and reuse of geographical data.

**Keywords:** Geographical Weighted Regression (GWR); Temporal-spatially varying impacts; COVID-19 Cumulative Case

## 1. Introduction

A new coronavirus, called coronavirus disease 2019 (COVID-19), is causing an outbreak of respiratory illness worldwide. It caused a deadly severe acute respiratory syndrome, WHO reported, initially found in Wuhan of China (Holshue et al., 2020; Moghadas et al., 2020; Sha et al., 2020; Yang et al., 2020). Given coronavirus disease (COVID-19) swept through the world, everything about people's mobilities is changed. COVID-19, as a global social, environmental, and economic comprehensive crisis, extremely impacts on people's daily life and reshape people's routine behaviors, especially a pervasive sense of quarantine fatigue and panic attacks of getting infected are challenging human's fortitudes (Ahmar et al.,2020;Bashir et al., 2020; Bilal et al, 2020). Most countries have been forced to take emergency measures, including closing cities, suspending school classes, restricting population movement, and keeping social distances, having great negative effects on economic development and resident's health (Jin et al., 2020; Yuan et al., 2020; P.D. et al, 2020; Menut et al., 2020). The U. S was called the first highest country with confirmed cases of COVID-19 in the world (Ahmar et al, 2020; Worldometer, 2020). Until Feb. 05, 2021, there are 27.3 million cumulative cases, 65,236 new cases, 468,780 total deaths cases, and 1,786 new deaths cases in the U.S. A severe economic downturn behind figures was predicted by considering how policy has supported businesses and families since March 2020. The first COVID-19 case in the United States was confirmed on January 19, 2020, in Washington State (Holshue et al., 2020; Ellis, et al, 2020; Qu, et al, 2020). From summer to winter in 2020, the virus, as a perfect storm, virtually spreads every part of the U.S. at a speed unprecedented in American history, according to Johns Hopkins University data. Jamie Ducharme argued the pandemic had claimed more than three times the American lives that were lost in the Vietnam War (Ducharme, 2020). The coast-to-coast surge is causing hospitals across the country to the edge of catastrophe. Doctors and

nurses exhausted and their intensive-care units running dangerously low on beds.

Texas, as the 28th state of the union in 1845, occupies the south-central segment of the country and is the largest state in area except for Alaska and California. The state extends nearly 1,000 miles (1,600 km) from north to south and about the same distance from east to west. Its average population growth has exceeded that of the country. Texas is facing issues associated with increased longevity and an aging population. The state's overall population is aging, and about one-tenth of Texans are over age 65. Considering the total number of coronavirus cases in the U.S, Texas plays a significant role in curbing the spread of COVID-19 in the U.S. Therefore, it is plausible to choose Texas as our study object of COVID-19. Ostensibly, the Texas government made stringent policy inventions to mitigate the spread of COVID-19 coronavirus based on unremitting Texas Executive Orders (TEO) and Public Health Disaster Declarations (PHDD). On March 16, 2020, the U.S promulgated the President's Coronavirus Guidelines for America, calling upon Americans to slow the spread of COVID-19 by avoiding social gatherings in groups of more than 10 people, using the drive-thru, pickup, or delivery options at restaurants and bars, and avoiding visitation at nursing homes, among other steps. Texas experienced 5 stages of COVID-19 risk-based guidelines, including the first phase on April 4th, the second phase on May 18th, the third phase on June 3rd, the fourth phase on July 4th, the fifth phase on Dec 15th for Travis County. Stage 5 signifies the most unfettered spread of the virus and includes the most stringent guidelines. In addition to curfew considerations, under Stage 5 guidelines, it is suggested that only essential businesses remain open and that everyone avoids non-essential travel and gatherings involving people outside of those in a person's immediate household. Those recommendations are in addition to social distancing measures such as avoiding the sick, wearing a facial covering, and maintaining distance from others. There are 28 executive orders regarding containing the spread of COVID-19 from Executive Order 8 on March 19 to Executive Order 32 on October 7. There are 11 times of publishing PHDD During March 19, 2020 -January 15, 2021. "The novel coronavirus (COVD-19) has been recognized globally as a contagious respiratory virus" mentioned on Gov. Abbott Issues Executive Order 8. On March 19, 2020, "Tests for human diagnostic purposes of COVID- 19 should submit to Texas Department of State Health Services (DSHS)." Claimed on March 24, 2020. The spread speed of COVID-19 is outpacing the density policy updating. Meanwhile, the once-in-a-100-year Nevertheless, coronavirus disease 2019 took a toll on Texas counties beyond words, from Facebook to Twitter, from nursing homes to children's daycare, from communities to churches, from groceries to restaurants, from elementary schools to universities (Sha et al. 2000). The socioeconomic impact of COVID-19 is well documented as well (Bashir et al, 2020 Sha et al, 2020, Yang et al, 2020). When tracing the COVID-19 footprint, the first case was announced by The Texas Department of State Health Services on March 4th in Fort Bend County. Texas surpasses 2,433,110 total COVID-19 cases and 128,000 deaths cases so far. These figures are dramatically increasing every single day. Some counties in Texas are already playing out their dystopian worst-case scenarios. In particular, the dead in El Paso have been shunted to mobile morgues partially staffed by the incarcerated (Hogue et al., 2020). Some of the critiques are thought that non-controlling the virus expansion is mainly responsible for potential policy from the Trump Administration, who has done little to counter rampant misinformation about the pandemic and has made numerous incorrect statements about the virus's origins, spread, and deadness. Indeed, it is worth noting that establishing a long-lasting, real-time, and dynamic emergency alert system of health will mitigate natural disasters and lower disaster risks. Accordingly, it is imperative to tease out the spatial-temporal changes of the COVID-19 pandemic spread, sensitive area of vulnerability, and most vulnerable groups based on the county level. The objective of this paper is to investigate spatial-temporal variability between population (age structure, race, gender) and COVID-19 cumulative cases at 254 Texas counties, in the context of considering of impacts of social-economic (unemployment, annual income) and environmental factors, via spatial stratified heterogeneity analysis of using GWR models. This

research will facilitate scientific recognition of COVID-19. In other words, if the COVID-19 data has a spatial-temporal resolution to capture the trajectories, both approaches are adequate for the spread of COVID-19 recognition (Camara, 2020).

Under global limelight, COVID-19 research is widespread and interdisciplinary concerning, triggering people's brainstorming, swept in the world. For example, social injustice during the pandemic is advanced by Sarah Blue et al. regarding asylum seekers im/mobility at the US–Mexico Border during the COVID-19 Pandemic (Blue et al.2020). There are 482958 academic journals published COVID-19 topic. In light of Web of Science, 567 articles are being published different journals, including health, religion, cultural studies, history philosophy of science, humanities multidisciplinary, and philosophy. COVID-19 in South Korea is investigated in policies and innovations (Doowon et al, 2020). COVID-19 pandemic and lockdown In India have led to the collapse of regular mental health services (Grover, 2020). Interestingly, a comparison about COVID-19 response between the UK, USA, Germany, and South Korea is rooted in the systemic weaknesses of neoliberalism. Timothy I. proposed that economic recession and austerity impeded healthcare investment in U.K and U.S (Timothy I. et al. 2020). Air quality impacts COVID-19 are mentioned in Europe and China (Menut et al. 2020; Liu et al., 2020). From a macro spectrum perspective, global collaboration and a data-driven systems approach will contribute to addressing the COVID-19 pandemic and potential public health crisis (Francisco et al.2020).

Since the COVID-19 preading represented geographical dependence, GIS can combine divergent spatial data sets based on georeferencing, promoting the integration of health data with contextual characteristics. At the same time, descriptive modeling research that depends on GIS strength has examined the spatial associations of COVID-19 with socioeconomic and environmental characteristics (Smith et al., 2020). Currently, the uneven distribution of the COVID-19 pandemic is well enough to represent patterns of spatial heterogeneity with GIS spatial tools, which incorporate multidimensional social, economic, and health consequences, exposing to geographical inequity and a long-term impact on global health accurately, no matter what linear-regressions or non-linear regressions (Rosenkrantz et al., 2020; Smith, 2020; Guliyev, 2020 ) For instance, Ansari Saleh Ahmar predicted COVID-19 confirmed cases in the U.S with SutteARIMA method (Ahmar et al. 2020). Yaowen Luo exhibited a spatial nonlinear analysis of the COVID-19 death rate in the U.S (Luo et al.2020). Unfortunately, GIS spatial-temporal analysis is seldom mentioned. Chaowei Yang et al. (2020) put forward a spatial-temporal COVID-19 paradigm through modeling socioeconomic patterns to analyze at a finer scale using weekly confirmed cases in Massachusetts. They merely took into account the poverty rate, educational attainment, elderly people rate, and income four variables (Yang et al.2020). The drawback is the lack of environmental variables to underpin the model. Although population mobility, age, and race, as significant factors, are mentioned in the research, Mollalo only considered black females infection risk of COVID as an explanatory variable, it is limited to get the outcome on the most vulnerable groups of COVID-19 (Mollalo 2020, Smith 2020, Lakhani 2020, Rosenkrantz et al., 2000). Different methods are used to observe the goodness-of-fit test of the regression (e.g., multiple geographical weighted regression method and geographical weighted random forest method, but they do not account for the significance of the single variable. Furthermore, previous study areas are based on the macro spectrum, involving China, India, Italy, and the U.S (Rakhohori et al., 2020, Mollalo et al., 2020; Xie et al., 2020), regional study of COVID is almost ignored. As community gathering is the main reason for COVID-19, micro-study is a dispensable part of COVID-19 research, which could provide useful implications for preventing the spread of the COVID-19 pandemic. Even if the spatial-temporal ontology and semantic COVID-19 are completely performed in the context of big data in the Age of AI (Sha et al. 2020), longitudinal detection and explanation of COVID-19 in the U.S is essential for dealing with the linear scenario in a local regression model. On the other hand, according to empirical study, it is important to select variables that reveal the degree of social vulnerability. This is because disparity of social vulnerability is determined by location-based variables, which incorporate

136 factors such as race/ethnicity and socioeconomic status, leading to encode the vulnerability to adverse health
137 outcomes such as negative effects of COVID-19.
138 Analysis of the relationship between these possible risk factors (e.g., AQI, race, gender) and COVID-19 in
139 different counties will help develop policies to prevent and control the spread of COVID-19 technically. The
140 spatial-temporal distribution of COVID-19 will also contribute to county-driven COVID-19 real-time and
141 dynamic monitor systems. The advantage is that the results are directly used to draw up community containment
142 strategies, which are fundamental public health measures used to control the spread of communicable diseases,
143 including isolation and quarantine (Mollalo et al., 2020). Therefore, this paper unveils spatial-temporal
144 heterogeneity on county-level within a state, providing real-time scientific evidence for creating an effective
145 disease surveillance system in COVID-19.

146 **2. Methodology**

147 2.1 Data source

148 Demographical factors such as age structure, gender, and race are examined to detect the most vulnerable
149 groups. Since PM2.5 and Air Quality Index (AQI) are used to measure the severity of air pollution, which is highly
150 related to respiratory diseases. PM2.5 and AQI are also considered explanatory variables. COVID-19 data
151 (cumulative cases and new cases) as dependent variables are from the Centers for Disease Control and Prevention
152 (CDC), COVID-19 fatality data based on death certificates. A fatality is counted as a COVID-19 fatality when the
153 medical certifier attests to the death certificate that COVID-19 is a cause of death. Mortality is equal to fatalities
154 divide by cumulative cases. Hospitalization (i.e., THB, BPC) from The Texas Department of State Health
155 Services (DSHS) is reported daily by hospitals through eight Hospital Preparedness Program providers that
156 coordinate health care system preparedness and response activities in Texas. They are viewed as explanatory
157 variables over 2020. Population data (e.g., race, diverse age group, gender, population density) as explanatory
158 variables are from the 2019 U.S Census Bureau. Economic data (e.g., annual income) as explanatory variables are
159 from the Texas Association of Counties, the statistical period is 2019. Environmental data is the result through
160 interpolating limited samples, which are from the United States Environmental Protection Agency (i.e., AQI,
161 PM2.5) and National Weather Service (i.e., temperature, precipitation), Statistical period is the whole year of 2020,
162 categorized four quarters. Quarterly data are real-time raw data at the end day of each quarter.

163 Table. 1 A List of Variables Used for Geostatistical Analysis

| Variable category | Variable name | Acronym | Variable description |
|---|---|---|---|
| Economic | Annual income | PCI | Annual income per 1,000 residents |
| | Unemployment | UEM | Percent of residents who don't have job |
| Environmental | Precipitation | PCN | Mean precipitation per month |
| | Temperature | TPE | Mean temperature per month |
| | PM2.5 | PM2.5 | Mean PM2.5 per day |
| | Air quality | AQI | Mean air quality per day |
| | Land Area | LA | Total land area per county |
| Demographic | Population density | POD | Population density |
| | Total population | TP | Total population |
| | Male population | PMP | Percent of residents who are male |
| | Black population | PBP | Percent of residents who are black |
| | Populationbetween20-59 | P59 | Percent of residents who are between 20-59 |
| | Population beyond 80 | P80 | Percent of residents who are beyond 80 |

| | | | |
|---|---|---|---|
| Health | Total hospital beds | THB | Total hospital beds |
| | Beds per capital | BPC | Incidents per 1,000 residents |
| Covid-19 | Cumulative Case | CC | Cumulative Case number |
| | New Case | NC | New case number per season |
| | Incidence Rate | IRP | Percent of Case on total population |
| Covid-19 | Fatalities | TF | Total death number |
| | Mortality Rate I | MR1 | Percent of fatalities case on total case |
| | Mortality Rate 2 | MR2 | Incidents per 10,000 residents |

## 2.2 Study Area

In this paper, 254 counties of Texas are our research objects.

## 2.3 Spatial-temporal Study Framework

For temporal study in the paper, time series thought was classified into four layers according to four seasons in 2020. Quarterly statistical data are based on environmental, social-economic indexes at the end of the season in response to COVID-19 NC and TC at that time. The temporal-study framework is in Fig. 1.

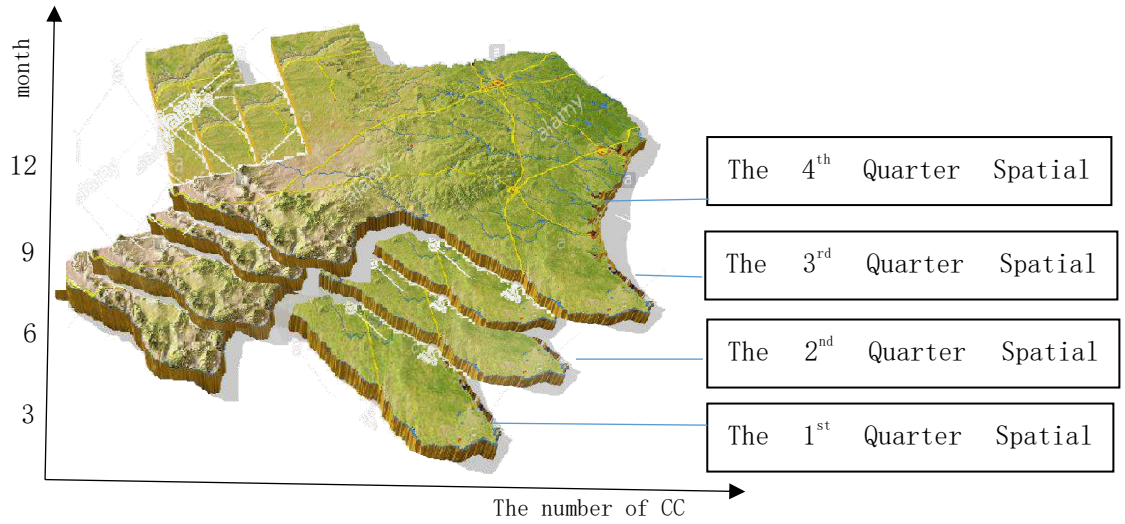



Fig. 1 Temporal-Study Framework

For spatial-study perspective, we explore correlations between variables with SPSS before building GWR models, no matter what kinds of variables. Since dependent variables must meet the assumption of a normal distribution, we have to describe their statistical characteristic property and spatial autocorrelation analysis. Simultaneously, all explanatory variables after standardization should be examined by Principal Component Analysis to eliminate multicollinearity. After that, we try to model simple Ordinary Lease Square (OLS) and geographically weighted regression between variables. Finally, via two models' comparisons, we pay more attention to their differences in spatial heterogeneity and analyze how did it happen, as shown in Fig.2.

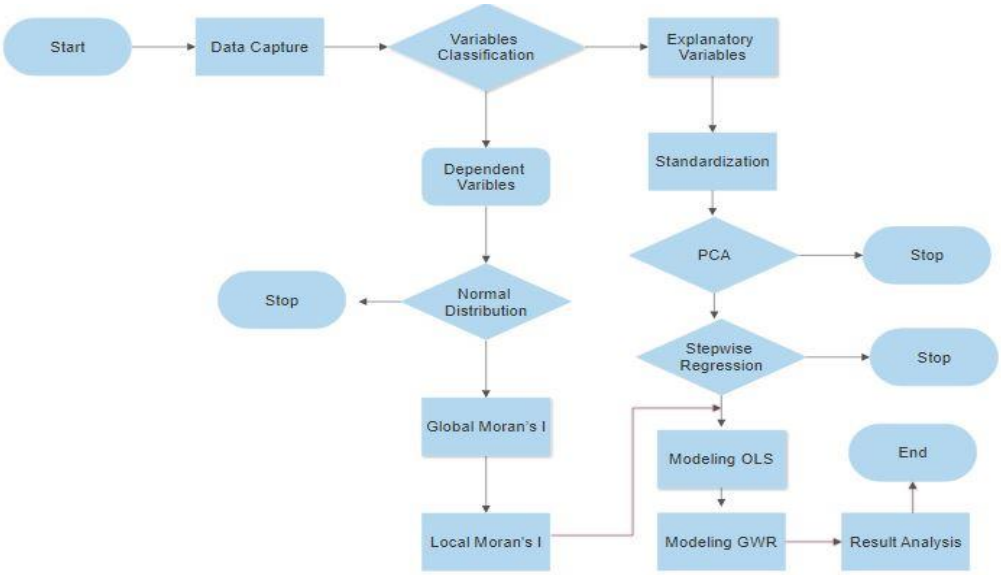



Fig. 2 Data Flow

## 2.4 data preprocessing and preparing

In Fig. 3, The number of NC and CC in the first season is the same as 3604. The number of NC in the second season is five times more likely than the first season. The number of NC in the third season is tetra more likely to the second season. The number of NC in the last season is double more likely to the third season. From the number of CC perspective, the number of CC is taking surge without turn points. That means the spread of COVID-19 is monotonically increasing without controlling.

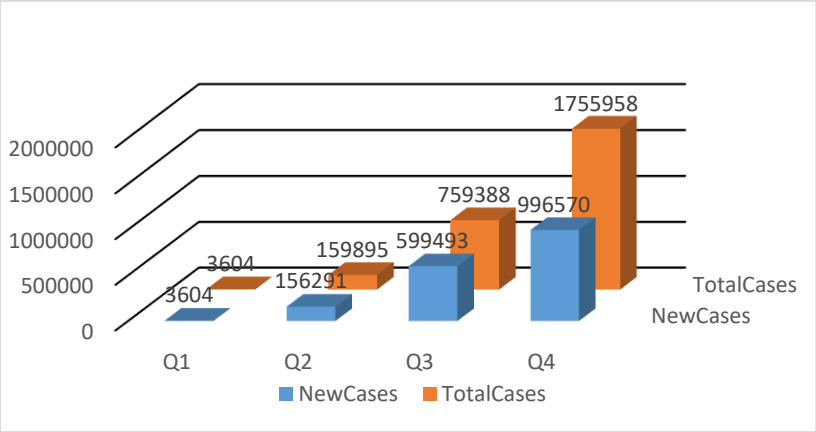



Fig.3 Texas Cases Changes over time in 2020

Data standardization is this process of making sure that your dataset can be compared to other data sets. It is a key part of the research, and standardized data is essential for accurate data analysis. It is also easier to make clear conclusions about current data when there are other data to measure it against. The condition of standardization with the Z-score is that the data mean is equal to 0 and the standard deviation is equal to 1.

To reduce the dimensionality of the dataset down to fewer explanatory variables, Principal Component Analysis (PCA) is one of the common techniques to avoid multilinearity without losing the attribution of variables. PCA could increase interpretability but at the same time minimize information loss. It does so by creating new

uncorrelated variables that successively maximize variance. In the PCA procedure, a set of possibly correlated variables is transformed into a set of linearly uncorrelated variables using the orthogonal transformation. This set of linearly uncorrelated variables is also called a PC. The number of PC extracted from PCA is less than or equal to the number of previous possibly correlated variables (Pai et al., 2017).

Stepwise Regression (SR) is an automatic variable selection procedure that selects from a couple of candidates the explanatory variables, which are the most related. We used the unidirectional forward methods. Forward selection begins with no variables in the model, examining each variable with a chosen model-fit criterion until none of the remaining variables improves the model to a statistically significant extent (Massimo, 2020).

## 3 Method

### 3.1 Ordinary Least Square

In regression analysis, Ordinary Least Squares (OLS) is a traditional method for estimating a linear regression between dependent variables and independent variables.

Simple OLS is the estimation of a linear relationship between two variables, $Yi,$ and $Xi$, of the form:

$$Y_i = \alpha + \beta \mathrm{X}_i + u_i \qquad \mathrm{i}=1, 2,\ldots\mathrm{n} \quad (1)$$

Where $Y_i$ denotes the ith observation on the dependent variable Y which could be CC, and Xi denotes the $i^{th}$ observation on the independent variable X which could be synthetic factors. OLS assumptions involve the disturbances have zero mean and a constant variance, in addition to are not correlated. The explanatory variable X in OLS is non-stochastic.

### 3.2 Geographical Weighted Regression

According to the first law of geography, there is more similarity between more adjacent geographical entities (Tobler, 1970). Meanwhile, due to the unbalanced distribution of natural resources endowment and socioeconomic factors in different provinces, there also exists interregional spatial correlation and spatial heterogeneity. And because of these, such global-regression-model-related assumptions do not hold anymore, for instance, data values are independent of geographical location, there exists no spatial correlation, and sample data are balanced. Therefore, it is impossible to properly explain an individual situation, and herein spatial heterogeneity, by using global overall parameters. Based on Foster's spatial varying parameter regression, a Geographically Weighted Regression model (GWR) (Fotheringham et al., 2002) was further proposed by Fotheringham using a local smooth processing method to solve the spatial heterogeneity. With spatial heterogeneity taken into consideration, geographic coordinates and core functions are utilized to carry out local regression estimation on adjacent individuals of each group. The equation of the GWR fitted model is in Eq(2) (Nakaya, 2016).

$$y_i = \beta_0(u_i,v_i) + \sum_k \beta_k(u_i,v_i)x_{k,i} + \varepsilon_i \quad (2)$$

where i denotes the individual sample; $(u_i,v_i)$ is the coordinates of sample i; $\beta_k(u_i,v_i)$ is the $k^{th}$ regression parameter of sample i; $y_i$ is the dependent variable of sample i , $x_{k,i}$ is the $k^{th}$ independent variable for the sample i, $\varepsilon_i$ is random error term which obeys normal distribution when the variance is a constant, thus the parameter estimation value of sample i is given by:

$$\hat{\beta}(u_i,v_i) = (X^T \mathrm{W}(u_i,v_i)X)^{-1}X^T W(u_i,v_i)y \qquad (3)$$

where W is the spatial weight matrix, whose selection and the setting is the core issue of GWR regression. And its calculation consists of two major steps. The first step is the selection of a proper kernel function to express a spatial relationship between the observed units. Specifically, four major kernel functions are being used in existing research, namely fixed Gaussian, fixed Bi-square, adaptive Bi-square, and adaptive Gaussia. Since the merits of a kernel function play a direct and decisive role in obtaining the most accurate possible regression parameter estimation of spatial heterogeneity, after careful analysis and comparison, Fixed Gaussian

was chosen as the kernel function in the paper, which is expressed as

$$w_{ij} = \exp(-d_{ij}^2/\theta^2) \quad (4)$$

where $w_{ij}$ represents the distance weight from sample i to sample j; $d_{ij}$ is the Euclidean distance between sample i and sample j; θ is the bandwidth, which determines the speed at which the spatial weight attenuates with distance. The second step of spatial weight matrix calculation is the selection of optimal bandwidth which could contribute to a higher fitting degree. According to the GWR4.09 User Manual (Nakaya, 2016), bandwidth selection criteria include AIC (Akaike Information Criterion), AICc (small sample bias-corrected AIC), BIC, and CV (Cross Validation).

## 4. **Results and Findings**

### 4.1 Normal Distribution

The precondition of regression analysis is that the dependent variable should meet the normal distribution. The request for normal distribution has two conditions. One is uncertain variable is symmetric about the mean, another is that uncertain variable is more likely to be in the vicinity of the mean than far away. Thus, a normal distribution is conducted in 5 dependent variables quarterly. After logarithm transformation, quarterly CC within is qualified normal distribution except for the first quarter in Fig.4. When modeling GWR regressions, the first quarter CC is overlooked as skewed distribution.

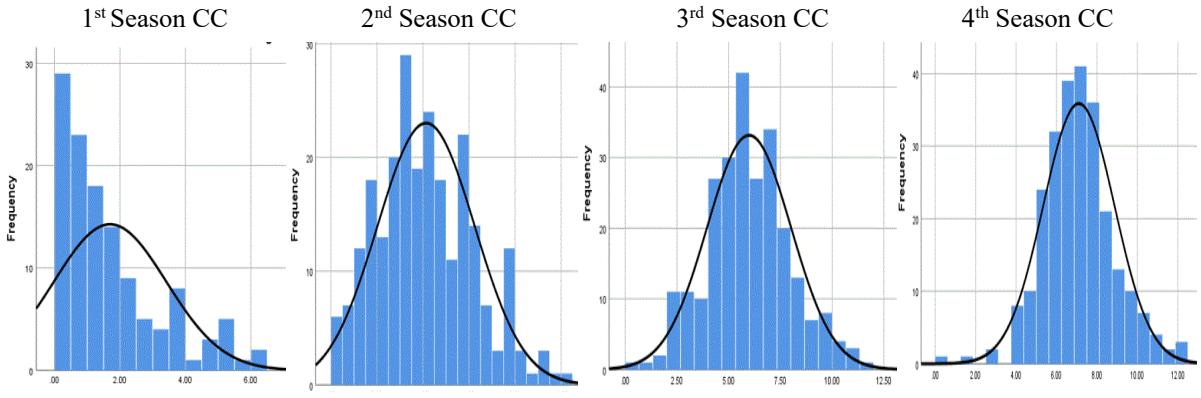



Fig. 4 CC distribution graph

### 4.2 Correlation

According to table 2, the first quarter CC is positively significant to THB, POD, PCI, TP, PBP, and P59. It is negatively significant to P80. The second quarter CC is positively significant to TPE, PCN, THB, POD, TP, PBP, UEM, P59, and negatively significant to P80. The third quarter CC is positively significant to TPE, PCN, AQI, THB, POD, TP, PBP, UEM, P59. While it is negatively significant to P80. The last quarter CC is positively significant to TPE, PCN, AQI, THB, POD, TP, PBP, UEM, P59, and negatively significant to P80. As the result, the correlation from the second quarter to the fourth quarter is similar.

Table 2 Person Correlation between CC and Explanatory Variable

| Explanatory Variables | CC Quarter 1 Coef. / Sig. | CC Quarter 2 Coef. / Sig. | CC Quarter 3 Coef. / Sig. | CC Quarter 4 Coef. / Sig. |
|---|---|---|---|---|
| TPE | 0.155/0.088 | 0.128*/0.042 | 0.365**/0.000 | 0.292**/0.000 |
| PCN | 0.038/0.679 | 0.307**/0.000 | 0.378**/0.000 | 0.325**/0.000 |
| AQI | 0.106/0.246 | 0.021/0.744 | 0.249**/0.000 | 0.260**/0.000 |
| THB | 0.645**/0.000 | 0.481**/0.000 | 0.495**/0.000 | 0.509**/0.000 |
| BPC | 0.154/0.091 | 0.047/0.454 | 0.036/0.573 | 0.097/0.123 |
| POD | 0.749**/0.000 | 0.570**/0.000 | 0.581**/0.000 | 0.600**/0.000 |
| LA | 0.133/0.145 | −0.430/0.499 | −0.066/0.297 | −0.031/0.628 |

| | | | | |
|---|---|---|---|---|
| PCI | 0.335**/0.000 | −0.020/0.753 | −0.048/0.450 | −0.024/0.702 |
| TP | 0.690**/0.000 | 0.512**/0.000 | 0.526**/0.000 | 0.541**/0.000 |
| PBP | 0.243**/0.007 | 0.455**/0.000 | 0.398**/0.000 | 0.362**/0.000 |
| UEM | −0.073/0.422 | 0.161**/0.010 | 0.181**/0.004 | 0.165**/0.008 |
| PMP | −0.174/0.056 | −0.055/0.380 | −0.053/0.398 | −0.077/0.219 |
| P59 | 0.467**/0.000 | 0.503**/0.000 | 0.474**/0.000 | 0.473**/0.000 |
| P80 | −0.451**/0.000 | −0.501**/0.000 | −0.450**/0.000 | −0.399**/0.000 |

Note: *Correlation is significant at the 0.05 level (2-tailed), ** Correlation is significant at the 0.01 level (2-tailed).

### 4.3 Spatial Autocorrelation

The Spatial Autocorrelation (Global Moran's I) tool measures spatial autocorrelation based on both feature locations and feature values simultaneously. Given a set of features and an associated attribute, it evaluates whether the pattern expressed is clustered, dispersed, or random. In this paper, Global Moran's I and local Moran's are implemented. The results of Global Moran of the accumulative case are more than 2.58, indicating quarterly CC is remarkable and clustered. In Anselin local Moran's I (Fig 5), the second season CC is classified four clusters, including HH, HL, LH, and LL. HH is distributed in northern Texas of 39 counties, HL is distributed at 11 counties, LH is limited at 6 counties, LL occupied the south and east of Texas of 41 counties. In the third season, HH is distributed in northern Texas of 44 counties, HL is distributed at 7 counties, LH is limited at 6 counties, LL is distributed south and eastern of Texas of 48 counties. In the fourth season, HH is distributed in northern Texas of 37 counties, HL is distributed at 7 counties, LH is limited at 7 counties, LL is distributed south and eastern of Texas of 31 counties.

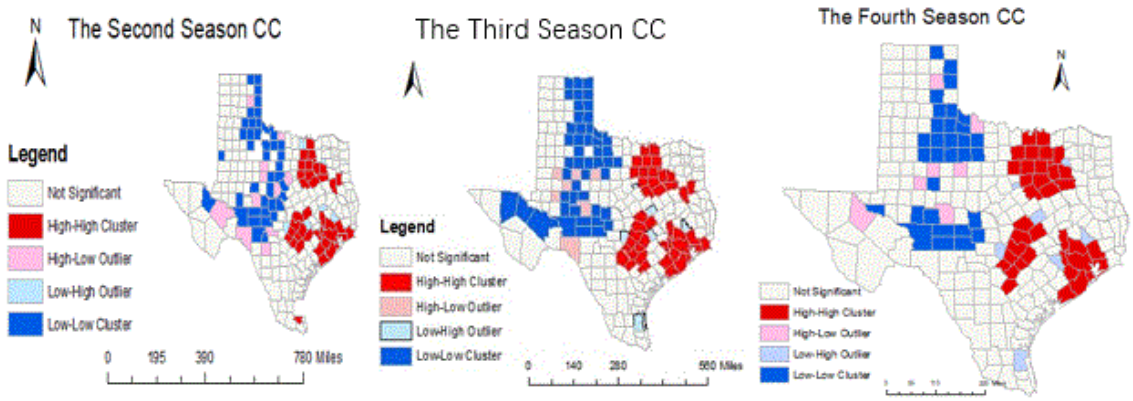



Fig. 5 Local Moran's Model of CC

### 4.4 Factor analysis

Through PCA, the dataset was examined using Kaiser-Meyer-Olkin (KMO) and Bartlett's Test of Sphericity. The KMO test compares the correlation statistics to identify if the variables include sufficient differences to extract unique factors. A KMO value of 0.616 for 14 explanatory variables is more than the threshold value of 0.5, The Bartlett's Test of Sphericity (BTS) value of 0.0 was significant ($p<0.001$), validating that correlation between variables does exist in the GWR models.

Communality is a common variance between 0 and 1, using the remaining variables as factors, was used to determine if any variables should be excluded from the factor analysis. 0.7 threshold is used to determine the significance of explanatory variables.

PCA was conducted as the factor analysis method within this paper. Given an eigenvalue threshold greater than 1.0, 6 components in the second quarter, 5 components in the third quarter and the fourth quarter are produced, explaining cumulative 76.57% (the second quarter), 70.60% (the third quarter), 70.81% (the fourth quarter) of the variance within the models. A varimax rotation was used to assist in the interpretation of the PCA analysis. The rotated component matrix was examined for variables with a cutoff threshold of 0.7. In the second quarter, the first factor exhibited high loading on variables related to THB, POD, and TP, indicating COVID-19 cases are positively related to hospitalization and total population. That means the population and medical care are two main indicators of COVID-19. Factor 2 was a composite age structure index related to P59 and P80, COVID-19 CC are positively related to 20-59 Population but negatively related to 80 Populations. That means 20-59 population directly contribute COVID-19 patients, P80 leads to CC reduction. That means race and weather are two underlying elements of COVID-19. Factor 3 is air quality, directly positive relevant to COVID-19 CC, indicating air quality improvement play positive roles in COVID-19 reduction. Factor 4 represents the economic index, including PCI and UEM. It represents annual income is negatively related to COVID-19 (i.e., a decrease of annual income, a high risk of COVID-19 infection), unemployment is positively related to COVID-19 (i.e., an increase of unemployment rate leads to a high risk of COVID-19 infection). Factor 5 is natural supply (i.e., LA), which is negatively related to COVID-19 CC. That demonstrated that keep spatial distancing benefits COVID-19 CC reduction. Factor 6 refers to medical supply (i.e., BPC), meaning hospital beds is positive related to COVID-19 CC. 5 factors in the third quarter is identical with the 5 factors in the first quarter except for factor 6. 5 factors in the last quarter are similar with the 5 factors in the first quarter except for factor 3. The distinction in the last quarter is that factor 3 is added to PCN in natural supply, meaning precipitation positively influences on COVID-19 increasing. The concreted relationships are shown in Table 3 and Table 4.

Table 3 Factors' roles in modeling OLS and GWR regressions

| No. | Items | Quarter 2 | Quarter 3 | Quarter 4 |
|---|---|---|---|---|
| 1 | Population and hospitalization | Factor1(THB, POD, TP) | Factor1(THB, POD, TP) | Factor1(THB, POD, TP) |
| 2 | Age structure | Factor2(P59, P80) | Factor2(P59, P80) | Factor4(P59, P80) |
| 3 | Air quality | Factor3(AQI) | Factor5(AQI) | |
| 4 | Economic | Factor4(PCI, UEM) | Factor4(PCI, UEM) | Factor2(PCI, UEM) |
| 5 | Natural supply | Factor5(LA) | Factor3(LA) | Factor3(PCN, LA) |
| 6 | Medical supply | Factor6(BPC) | | Factor5(BPC) |

Table 4 Rotated Component Matrix

| Variables | The second quarter component | | | | | | | The Third quarter component | | | | | | The Fourth quarter component | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Extract | 1 | 2 | 3 | 4 | 5 | 6 | Extract | 1 | 2 | 3 | 4 | 5 | Extract | 1 | 2 | 3 | 4 | 5 |
| TPE | 0.771 | 0.140 | -0.001 | 0.826 | 0.230 | 0.131 | -0.012 | 0.795 | 0.142 | -0.063 | 0.48 | 0.32 | 0.662 | 0.616 | 0.145 | 0.548 | 0.213 | 0.027 | -0.498 |
| PCN | 0.698 | 0.115 | 0.097 | -0.454 | 0.307 | 0.582 | -0.190 | 0.81 | 0.122 | -0.118 | 0.647 | 0.368 | 0.476 | 0.769 | 0.104 | 0.372 | **0.769** | 0.056 | -0.163 |
| AQI | 0.742 | -0.023 | -0.023 | 0.828 | -0.006 | -0.181 | -0.149 | 0.633 | 0.294 | 0.021 | -0.191 | 0.031 | **0.713** | 0.504 | 0.175 | 0.551 | 0.24 | -0.022 | -0.334 |
| THB | 0.956 | 0.973 | 0.083 | 0.021 | -0.022 | 0.003 | 0.043 | 0.952 | **0.974** | 0.056 | 0.013 | -0.015 | -0.006 | 0.958 | **0.974** | 0.012 | 0.024 | 0.082 | 0.044 |
| BPC | 0.779 | 0.108 | -0.099 | -0.090 | -0.010 | 0.070 | 0.863 | 0.343 | 0.148 | 0.04 | 0.048 | 0.081 | -0.558 | 0.642 | 0.121 | -0.003 | 0.085 | -0.053 | **0.786** |
| POD | 0.927 | 0.942 | 0.158 | 0.033 | -0.053 | 0.097 | -0.020 | 0.926 | **0.945** | 0.124 | 0.1 | -0.056 | 0.072 | 0.926 | **0.942** | -0.013 | 0.115 | 0.151 | -0.039 |
| LA | 0.755 | 0.083 | 0.072 | -0.111 | 0.198 | -0.831 | -0.018 | 0.71 | 0.069 | 0.062 | **-0.804** | 0.18 | 0.148 | 0.613 | 0.077 | 0.195 | **-0.747** | 0.1 | 0.034 |
| PCI | 0.658 | 0.145 | 0.076 | -0.015 | -0.782 | 0.109 | -0.083 | 0.69 | 0.138 | 0.046 | 0.097 | **-0.807** | 0.094 | 0.626 | 0.177 | **-0.728** | 0.093 | 0.071 | -0.227 |
| TP | 0.972 | 0.978 | 0.117 | 0.038 | -0.029 | 0.008 | -0.012 | 0.967 | **0.978** | 0.08 | 0.019 | -0.028 | 0.061 | 0.972 | **0.979** | 0.013 | 0.031 | 0.112 | -0.02 |
| PBP | 0.683 | 0.280 | 0.229 | -0.223 | 0.349 | 0.603 | 0.133 | 0.593 | 0.293 | 0.274 | 0.514 | 0.314 | -0.262 | 0.71 | 0.227 | 0.258 | 0.698 | 0.234 | 0.223 |
| UEM | 0.721 | 0.032 | 0.019 | 0.159 | 0.815 | 0.129 | -0.115 | 0.676 | 0.025 | 0.004 | 0.126 | **0.801** | 0.136 | 0.658 | 0 | **0.806** | 0.071 | 0.014 | -0.062 |
| PMP | 0.529 | -0.169 | 0.430 | -0.046 | -0.016 | -0.111 | 0.549 | 0.358 | -0.141 | 0.533 | -0.161 | 0.013 | -0.168 | 0.454 | -0.159 | -0.019 | -0.077 | 0.46 | 0.459 |
| P59 | 0.796 | 0.162 | 0.851 | 0.064 | -0.094 | 0.172 | 0.053 | 0.786 | 0.185 | **0.835** | 0.196 | -0.101 | 0.075 | 0.777 | 0.17 | -0.077 | 0.17 | **0.844** | -0.034 |
| P80 | 0.700 | -0.189 | -0.805 | 0.087 | -0.047 | 0.065 | 0.041 | 0.646 | -0.209 | **-0.774** | 0.045 | -0.025 | -0.022 | 0.688 | -0.188 | -0.034 | 0.052 | **-0.805** | 0.028 |

4.5 Comparison of composite OLS & GWR models

Modeling OLS is to examine whether there is a linear relationship between CC and its factors. By the T-test and F test, all factors are passed. Modeling GWR is to examine whether there is a spatial-temporal relationship between CC and its factors. Since COVID-19 CC is clustered and varies around the study area, and ADAPTIVE kernel in GWR models is appropriate. The AICc method I chose to find the bandwidth which minimizes the AICc value – the AICc is the corrected Akaike Information Criterion (it has a correction for small sample sizes). In the second quarter, the AICc value is decreased from 883.74 in the OLS model to 811.99 in the GWR, $R^2$ is changed from 0.54 in the OLS model to 0.77 in the GWR model. Neighbors are declined from 254 neighbors in the OLS models to 77 neighbors in the GWR models. In the third quarter, the AICc value is decreased from 870.29 in the OLS model to 790.31 in the GWR, $R^2$ is changed from 0.55 in the OLS model to 0.77 in the GWR model. Neighbors are declined from 254 neighbors in the OLS models to 77 neighbors in the GWR models. In the fourth quarter, the AICc value is decreased from 850.42 in the OLS model to 778.75 in the GWR, $R^2$ is changed from 0.49 in the OLS model to 0.72 in the GWR model. Neighbors are declined from 254 neighbors in the OLS models to 83 neighbors in the GWR models. All residuals of the GWR maps are lower, less than that of the OLS maps. Predicted CC in GWR quarterly map is more clustered than OLS quarterly map, the cluster area is in eastern and northern Texas. Therefore, the GWR model is superior to the OLS model. In table 5,

Table 5 GWR & OLS models' Comparison List

| items | quarterly | 5fOLS | 5fGWR |
|---|---|---|---|
| AICc | 2 | 883.74 | 811.99 |
| $R^2$ | 2 | 0.54 | 0.77 |
| Std. Deviation | 2 | 1.55 | 1.73 |
| Neighbors | 2 | 254 | 77 |
| Max_Value | 2 | 14.58 | 10.78 |
| Min_Value | 2 | 0.37 | 0.74 |
| Sum | 2 | 1008.23 | 1020.12 |
| Average | 2 | 4.13 | 4.18 |
| AICc | 3 | 870.29 | 790.31 |
| R2 | 3 | 0.55 | 0.77 |
| Std. Deviation | 3 | 1.464 | 1.626 |
| Neighbors | 3 | 254 | 77 |
| Max_Value | 3 | 15.589 | 12.31 |
| Min_Value | 3 | 3.078 | 1.595 |
| Sum | 3 | 1500.37 | 1505.32 |
| Average | 3 | 5.954 | 5.973 |
| AICc | 4 | 850.42 | 778.75 |
| R2 | 4 | 0.49 | 0.72 |
| Std. Deviation | 4 | 1.24 | 1.42 |
| Neighbors | 4 | 254 | 83 |
| Max_Value | 4 | 16.49 | 12.94 |
| Min_Value | 4 | 4.57 | 3.64 |
| Sum | 4 | 1804.85 | 1809.34 |
| Average | 4 | 7.11 | 7.12 |

336

337 4.6 GWR result analysis

338 4.6.1 Spatial change of CC factors.

339 Based on existing research, COVID-19 quarterly GWR models are also implemented in the research eara(Lau
340 et al. 2020, Mollalo et al. 2020). Fig. 4 incorporates Texas spatial-temporal distribution maps based on 6 factors
341 in terms of 6 aspects in table 3 in three quarters.

342 In the second quarter, factor 1 among 6 factors has the largest effects on CC in northern Texas thanks to the
343 maximum coefficient is 6.88. It's the lowest impact in eastern Texas due to the coefficient range of 0.61-0.88.
344 indicating total population and hospitalization are the key factor of COVID-19 and northern Texas is the main
345 precaution and control area of COVID-19. Factor 2 (Age structure) positively affects COVID-19 spatial
346 heterogeneity in central Texas with pink color. The area of the largest coefficient range 1.41-1.6 is distributed in
347 northern Texas. The smallest impacts of the coefficient range 0.43-0.83 are the coastal area at the bottom of the
348 map. Factor 3 is an air quality index, having remarkable spatial disparity for its coefficient is from range -1.78--
349 1.09 to range 0.15-0.66. In central Texas, the improvement of air quality is driven by COVID-19 CC, but it
350 reversely works in northern Texas. That indicates AQI has spatial non-stationary and environmental harness is
351 available reducing CC in northern Texas. Factor 4 is an economic composite index that coefficient is from range
352 -0.86—0.50 to range 0.56-0.81. The spatial heterogeneity is located between northern Texas, coastal counties, and
353 eastern Texas. Factor 5 is the natural supply index that coefficient is from range 0.02-0.28 to range 1.24-1.90. The
354 spatial heterogeneity is subtle. Factor 6 is the medical supply index that coefficient is from range -0.56--0.31 to
355 range 0.32-0.51. It is evident to see the change of spatial heterogeneity that the medical condition in northern
356 Texas is worse than in other Texas counties (https://www.dallasnews.com/news/2021/01/22).

357 In the third quarter, factor 1 among 6 factors is the dominant effect on CC due to the maximum range of
358 coefficient is 4.86-7.22. It's the lowest impact in central Texas due to the coefficient range of 0.63-1.00, implying
359 it is the most important factor and the spatial heterogeneity is subtle. Factor 2 (Age structure) positively affects
360 COVID-19 spatial heterogeneity in the eastern area with pink color. The area of the largest coefficient range 1.05-
361 1.22 is distributed in northern Texas. The smallest impacts of the coefficient range 0.18-0.45 are on 15 counties
362 of western Texas. Factor 3 is a natural supply index, having remarkable spatial disparity for its coefficient is from
363 range -1.10—0.26 to range 0.83-1.36. In central Texas, land area is driven COVID-19 CC, but it reversely works
364 on northern Texas. That indicates spatial distancing is not available for northern Texas, compared to central Texas.
365 Factor 4 is an economic composite index that coefficient is from range -0.49--0.28 to range 0.54-0.82. The spatial
366 heterogeneity is located between Central Texas, coastal counties, and eastern Texas. Factor 5 is air quality index
367 that coefficient is from range -1.09--0.41 to range 0.84-1.34. The spatial heterogeneity is obvious to be seen in the
368 change of spatial heterogeneity that positive impacts are from western Texas to eastern Texas while negative
369 impacts are from north Texas to western and southern Texas.

370 In the fourth quarter, factor 1 among 6 factors is still the dominant effect on CC without the range of
371 maximum coefficient is 3.99-6.6. Spatial heterogeneity is slight, implying it is a fixed factor. Factor 2 is an
372 economic composite index that coefficient is from range -0.58--0.23 to range 1.12-1.74. The spatial heterogeneity
373 is that areas of positive impacts are decreased while areas of negative impacts are increased. Factor 3 is a natural
374 supply index that coefficient is from range -0.28-0.03 to range 1.65-2.49. The spatial heterogeneity is that areas
375 of positive impacts are decreased while areas of negative impacts are moved from north Texas to eastern Texas.
376 Factor 4 is Age structure index that coefficient is moved from range 0.28-0.49 to range 1.02-1.16. The spatial
377 heterogeneity is that both areas of positive impacts and negative impacts are increased. Factor 5 is air quality
378 index that coefficient is from range -1.09--0.41 to range 0.84-1.34. The spatial heterogeneity is obvious to be seen
379 in the change of spatial heterogeneity that positive impacts are from western Texas to eastern Texas while negative

impacts are from north Texas to western and southern Texas. Factor 6 is the medical supply index that coefficient is from the range --0.87--0.52 to range 0.29-0.58. It is evident to see the change of spatial heterogeneity that areas of positive impacts are moved from eastern Texas to western and south Texas while areas of negative effects are decreased and moved.

4.6.2 Temporal change of CC factors

Population and hospitalization impact on COVID-19 within 3 quarters is relatively positive without a change in terms of two aspects. For coefficients, the value of the coefficient is fixed between 0.52 and 7.22. For the movement of spatial impacts, the spatial distribution of COVID-19 impacts is stagnant across three quarters. Importantly, northern Texas, including El Paso, Odessa, Midland, Lubbock, and Amarillo areas, are the most important areas of curbing Texas COVID-19 CC spread. Hence, community containment measures are the crucial result of cluster spreading as one of the characteristics of COVID deterioration.

Age structure impacts during 3 quarters are stably positive regarding two aspects. First, the coefficients from the second quarter to the fourth quarter are still, accounting for 0.28-1.60. Second, the spatial distribution of COVID-19 impacts is increased across three quarters. The areas of positive impacts with red color are sprawling while the areas of small impacts with blue color are extending. That means policy restrictions are gradually working and the virus is extremely spreading along with geographical trajectory.

Air quality impacts during three quarters are flexible in terms of two aspects. First, the coefficient range in two quarters is increased from -1.78-0.66 to -1.09-1.34. It demonstrated that the role of environments is rising. Second, both the areas of positive impacts with red colors and the areas of negative impacts with blue colors are moved from south to north, from north to west, respectively. Interestingly, air quality impacts are ignored in the fourth quarter, compared to other quarters. It implies that the rules of environmental impacts are a complicated and stochastic process.

Economic impacts during three quarters are flexible as well. On one hand, the coefficient range in three quarters is increased from -0.86-0.81 to -0.58-1.74. It demonstrated that the role of economic impacts is rising. Second, the areas of positive impacts with red colors are decreasing surround north Texas, whereas the areas of negative impacts with blue colors are extending around Houston. Interestingly, coastal county of positive impacts in the bottom of the map is shrinking until it is disappeared in the last quarter. It reveals that policy controlling and human self-consciousness are beneficial for mitigating COVID-19 spread.

Natural supply impacts in three quarters have fluctuated. First, the coefficient range within three quarters are changed 0.02-1.9, -1.10-1.36, into -0.28-1.49. It demonstrated that the role of natural supply is out of control. Second, the cluster of positive impacts with red colors is decreasing from 24 counties in north Texas to 9 counties. Simultaneously, the areas of negative impacts with blue colors are changing from the east to the north, finally landing on the east. It means that natural impacts are weakening, compared to other factor's impacts.

Medical supply impacts in three quarters have fluctuated as well. First, the coefficient range within two quarters is changed -0.5-0.51, 0, into -0.87-0.58. It demonstrated that the role of medical supply impacts is slight and out of control. Second, the cluster of positive impacts with red colors is increasing from the east-south tracking to the west-south tracking. Simultaneously, the areas of negative impacts with blue colors are decreasing from the center to the north. Interestingly, the impacts of the third quarter are ignored, representing medical capacity is limited and scarce.

## 5. Discussion

In this study, 14 potential risk variables are selected from the race, climate, land cover, demographic categories, hospitalization, gender, age structure, and socioeconomic as independent variables to estimate their spatial-temporal impacts on the distribution of the COVID-19 cumulative cases at the county-level in Texas. Since current research is lack of consideration of time series models, spatial-temporal GWR is explored to accurately identify

the imbalanced distribution of COVID-19 cumulative cases and the complex relationship between the COVID-19 CC and its risk factors (Luo et al. 2020). Four quarters in 2020 are categories to model quarterly GWR model to observe COVID-19 CC temporal-spatial change in Texas county. A spatial-temporal COVID-19 trajectory is simulated in the aforementioned analysis. Population, hospitalization, and age structure are exhibited stable, positive influences on COVID-19 cumulative cases. Climate, natural, economic, and medical conditions are displayed non-stationary, stochastic change processes. The longitudinal monitor mechanism bridges the gap of geographical analysis of COVID-19. Spatial-temporal geographical analysis is the main part of Spatial-Temporal Information System (STIs), which is defined by the positions of objects within the environment, the use of dynamic time intervals, ontology or the study of the relationships of the objects, real-time or real-world modeling, and the use of analytical tools. It is a mix of conventional Geographical Information Systems (GIS) with the use of modeling and simulation skills (Mcneil et al.2013).

Previous studies have shown that many social-environmental and economic variables are captured to analyze the distribution of COVID-19 cases, death rate with GIS tools via multiple (Desjardins et al. 2020; Shim et al. 2020; Lau et al. 2020, Mollalo et al. 2020) patterns of spatial change such as health status, disaster, transportation, atmosphere, climate, and socioeconomic indices, though they did not mention dimension reduction methods to avoid multicollinearity. In this study, for purposes of demonstrating the effectiveness of environmental and social-economic contributions on COVID-19 CC, PCA is used to simply multiple dimensions in CC spatial-temporal heterogeneity research. It is a useful statistical technique that has found application in fields such as face recognition and image compression and is a common technique for finding patterns in data of high dimension. The principal components of the faces in the training set are calculated. Recognition is achieved using the projection of the face into the space formed by the eigenfaces (Zou et al. 2018). Especially, composite factors are optimized GWR models fitness so that the models are catered for the demand of reality.

As previously described, variables are weighted to reflect their relative significance and relationship with other variables. However, researchers ought to figure out the meaning of synthetic factors. In this study, the findings are included: (1). The most important quarantine areas of COVID-19 in Texas are El Paso, Odessa, Midland, Lubbock, and Amarillo areas. (2). 20-59 population is the main source of Cumulative Case with a lower death rate, while over 80 population have lower infection rates and higher COVID death rates. Thus, over 80 population is the most vulnerable group of COVID-19. (3). Race and gender should be paid no attention to controlling COVID-19 since they are not components of factors. (4). Economic, environmental, race, and natural condition factors directly facilitated COVID-19 cumulative cases change with spatial-temporal heterogeneity.

This study is helpful to reshape disease surveillance and response systems, which remain the core of modern public health practices. COVID-19 extremely challenged the critical role of surveillance systems in offering timely and reliable health information to inform operational and strategic decision-making for multi-level health systems (Gadicheria et al. 2020). An ideal integrated disease surveillance and response system should collect and transmit data in real-time to all the stakeholders. It should be able to incorporate data from existing surveillance systems in real-time and analyze data to devise rapid response strategies (Bashir et al. 2020). Spatial-temporal Geostatistical Analysis on COVID-19 CC is selected real-time raw to expose real scenarios in Texas counties. Accordingly, the results of this research will provide evidence for current disease surveillance and response systems to facilitate the appropriate uptake and reuse of geographical data.

Spatial-temporal Geostatistical Analysis on COVID-19 CC can also be employed for local government and health care organizations to make scientific judgments of Covid-19 expansion and target vulnerable communities in Texas to prevent person-to-person spreading. Besides, it benefits to adjust health care resource

distribution in pandemic in Texas.

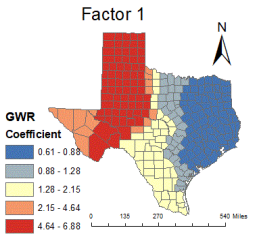
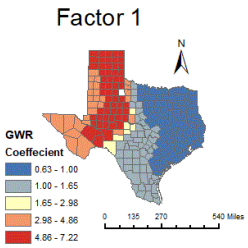
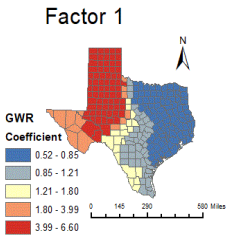
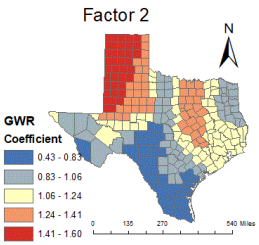
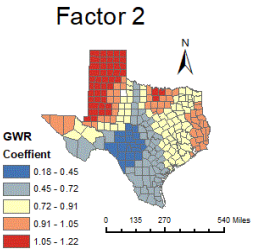
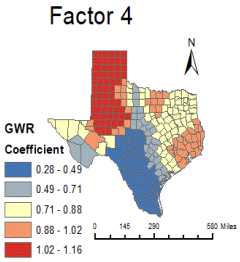
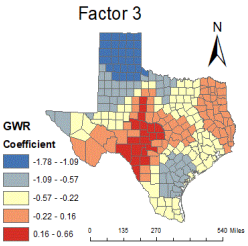
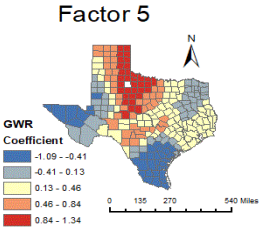
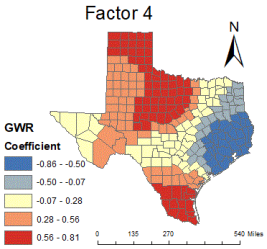
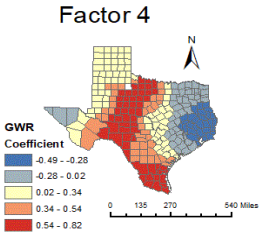
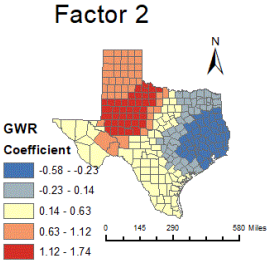
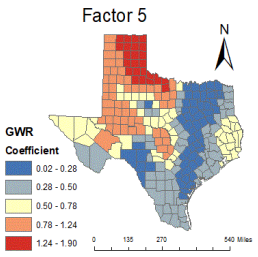
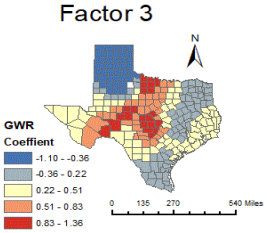
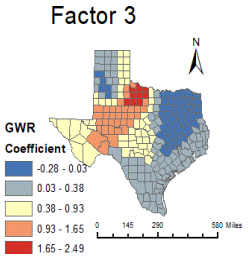
| No. | The 2nd Quarter | The 3rd Quarter | The 4th Quarter |
|---|---|---|---|
| 1 |  |  |  |
| 2 |  |  |  |
| 3 |  |  | |
| 4 |  |  |  |
| 5 |  |  |  |


Fig. 4 Factors spatial-temporal distribution of CC in the GWR Model

## 6 Conclusion

### 6.1 Limitations

This research just focuses on the Texas Covid-19 scenario, the application of research is limited to apply other states. There is no chronic disease data to supports this research. As explanatory variables, they should be incorporated in future studies, although are excited to see diabetes (Gupta et al., 2020) and cardiovascular conditions (Du et al., 2020) potential impacts on COVID-19 health outcomes. Collecting data of multiple dimensions might improve and enrich spatial variability findings of COVID-19. This research merely intended to spatial-temporal quarterly GWR models, yet there is a distance to be reached for real-time dynamic GWR models. GTWR or more effective spatial-temporal models are considered in future research.

### 6.2 Implications

The COVID-19 pandemic revealed systemic flaws in the food distribution system that fails to protect against hunger and dietinfluenced non-communicable diseases. It also exposes the conditions that made people who are living on low-incomes, disenfranchised, discriminated-against, and chronically ill the most vulnerable to harm from COVID-19. Especially COVID-19 accelerated the decline of health in the USA. Under the Trump administration's healthharming policies, some state and local governments have stepped in to protect their residents, but have done so unevenly. Some have attempted to maintain environmental and health insurance regulations, fund health insurance expansions, and protect immigrants. However, other local authorities have done the opposite by imposing work requirements on Medicaid enrollees, restricting access to abortion and contraception, and collaborating with federal agencies in apprehending and detaining immigrants. This geographical policy divergence continues a trend of devolution of responsibility for regulation and social service provision from federal to state and local levels (Woolhandler et al., 2020). This research will benefit geographical health divides evenly and provide food nutrition distribution reference transparently. Inspired by Mollalo et al. (2020) and Luo et al. (2020), who applied and compared the performance of multiscale GWR models across the United States for incident rates and death rates to account for the spatial variability of COVID-19, spatial-temporal GWR models are considered to compare global of OLS model to disclosure different change of COVID-19 cumulative case in response to social-economic and environmental variables at county-level in Texas. To add spatial-temporal variability understanding of empirical COVID-19 analysis, there is a lack of county-level research on COVID-19 GWR modeling. Therefore, the results of this study provide new empirical evidence to support future geographic modeling of the diseases.

**Author contributions** Xiu Wu: conceptualization, writing—original draft, formal analysis, methodology, software, formal analysis, investigation, visualization. F. Benjamin Zhan and Edwin T Chow: supervision and project administration, Jinting Zhang: conceptualization, validation, writing— review and editing, supervision.

**Funding** This research does not have any support.

**Availability of data and materials**: The datasets used during the current study are available from the corresponding author on reasonable request.

**Compliance with ethical standards**

**Competing interests**

The authors declare that they have no competing interests.

**Ethical approval** Not applicable.

**Consent to publish** All the co-authors consent the publication of this work.

**Consent to participate** Not applicable.

## Reference


Ahmar, Ansari Saleh, and Eva Boj. 2020. "Will COVID-19 Confirmed Cases in the USA Reach 3 Million? A Forecasting Approach by Using SutteARIMA Method." *Current Research in Behavioral Sciences 1* (November). doi:10.1016/j.crbeha.2020.100002.

Bag, R., Ghosh, M., Biswas, B., & Chatterjee, M. (2020). Understanding the spatio-temporal pattern of COVID-19 outbreak in India using GIS and India's response in managing the pandemic. *Regional Science Policy & Practice, 12(6), 1063.*

Bashir, Muhammad Farhan, Benjiang MA, and Luqman Shahzad. 2020. "A Brief Review of Socio-Economic and Environmental Impact of Covid-19." *Air Quality, Atmosphere & Health: An International Journal* 13 (12): 1403. doi:10.1007/s11869-020-00894-8.

Bashir A, Malik AW, Rahman AU, Iqbal S, Cleary PR, Ikram A. MedCloud: Cloud-Based Disease Surveillance and Information Management System. *IEEE Access, Access, IEEE*. 2020;8:81271-81282. doi:10.1109/ACCESS.2020.2990967

Bilal, Faiza Latif, Muhammad Farhan Bashir, Bushra Komal, and Duojiao Tan. "Role of Electronic Media in Mitigating the Psychological Impacts of Novel Coronavirus (COVID-19)." *Psychiatry Research 289* (July 2020): 113041. doi: 10.1016/j.psychres.2020.113041.

Cartenì, A., Di Francesco, L., & Martino, M. (2021). The role of transport accessibility within the spread of the Coronavirus pandemic in Italy. *Safety Science, 133.* https://doi-org.libproxy.txstate.edu/10.1016/j.ssci.2020.104999

Cássaro, F. A. M., & Pires, L. F. (2020). Can we predict the occurrence of COVID-19 cases? Considerations using a simple model of growth. *Science of the Total Environment, 728.* https://doi-org.libproxy.txstate.edu/10.1016/j.scitotenv.2020.138834

Congressional Research Service, 2020. Global Economic Effects of COVID-19. Retrieved from. https://fas.org/sgp/crs/row/R46270.pdf.

Dexuan Sha, Anusha Srirenganathan Malarvizhi, Qian Liu, Yifei Tian, You Zhou, Shiyang Ruan, Rui Dong, et al. 2020. "A State-Level Socioeconomic Data Collection of the United States for COVID-19 Research." *Data* 5 (118): 118. doi:10.3390/data5040118.

Development of a Composite Model of Quality of Life: A Case Study in Austin, Texas. (2012). *GIScience & Remote Sensing*, *49*(6), 802–821. https://doi-org.libproxy.txstate.edu/10.2747/1548-1603.49.6.802

Du, H., Wang, D. W., & Chen, C. (2020). The potential effects of DPP-4 inhibitors on cardiovascular system in COVID-19 patients. *Journal of Cellular & Molecular Medicine, 24(18), 10274–10278.* https://doi-org.libproxy.txstate.edu/10.1111/jcmm.15674

Ducharme, Jamie. "Class of COVID-19." *TIME Magazine*, vol. 197, no. 1/2, Jan. 2021, pp. 38–

43. EBSCOhost, search.ebscohost.com/login.aspx?direct=true&db=a9h&AN=147960785&site=eds-live&scope=site.

Ellis, Robert Evan. 2020. COVID-19: Shaping a Sicker, Poorer, More Violent, and Unstable Western Hemisphere. Strategic Studies Institute, United States Army War College. https://search-ebscohost.com.libproxy.txstate.edu/login.aspx?direct=true&db=cat00022a&AN=txi.b5452419&site=eds-live&scope=site

Fotheringham, A.S., Charlton. M.E., 2002. Geographically Weighted Regression: the Analysis of Spatially Varying Relationships. Wiley. New York. Chris Brunsdon.

Gray, V. (2017). *Principal Component Analysis: Methods, Applications, and Technology*. Nova Science Publishers, Inc.

Gilberto Câmara. On the semantics of big Earth observation data for land classification. *Journal of Spatial Information Science.* 2020;2020(20):21-34. doi:10.5311/JOSIS.2020.20.645

Guidolin, M., & Pedio, M. (2020). Forecasting commodity futures returns with stepwise regressions: Do commodity-specific factors help? *Annals of Operations Research,* https://doi-org.libproxy.txstate.edu/10.1007/s10479-020-03515-w

J. Harcourt, A. Tamin, X. Lu, S. Kamili, S.K. Sakthivel, J. Murray, *et al.*Isolation and characterization of SARS-CoV-2 from the first US COVID-19 patient *BioRxiv (2020),* 10.1101/2020.03.02.972935

Lakhani, A. (2020). Which Melbourne Metropolitan Areas Are Vulnerable to COVID-19 Based on Age, Disability, and Access to Health Services? Using Spatial Analysis to Identify Service Gaps and Inform Delivery. *Journal of Pain and Symptom Management*, *60*(1), e41–e44. https://doi-org.libproxy.txstate.edu/10.1016/j.jpainsymman.2020.03.041

Liu, Q., Sha, D., Liu, W., Houser, P., Zhang, L., Hou, R., Lan, H., Flynn, C., Lu, M., Hu, T., & Yang, C. (2020). Spatiotemporal Patterns of COVID-19 Impact on Human Activities and Environment in Mainland China Using Nighttime Light and Air Quality data. *Remote Sensing*, 12(10), 1576. https://doi-org.libproxy.txstate.edu/10.3390/rs12101576

Luo, Y., Yan, J., & McClure, S. (2020). Distribution of the environmental and socioeconomic risk factors on COVID-19 death rate across continental USA: a spatial nonlinear analysis. *Environmental Science and Pollution Research,* https://doi-org.libproxy.txstate.edu/10.1007/s11356-020-10962-2

Mollalo, A., Vahedi, B., & Rivera, K. M. (2020). GIS-based spatial modeling of COVID-19 incidence rate in the continental United States. *Science of the Total Environment, 728.* https://doiorg.libproxy.txstate.edu/10.1016/j.scitotenv.

Nakaya, T., (2016). GWR4.09 User Manual. Pp. 2-27.

Qu, Jie-Ming, Bin Cao, and Rong-Chang Chen. 2020. *Covid-19. [Electronic Resource]: The Essentials of Prevention and Treatment.* Elsevier. https://search-ebscohost-com/login.aspx?direct=true&db=cat00022a&AN=txi.b5571489&site=eds-live&scope=site.

Rosenkrantz, L. (1), Schuurman, N. (1), Bell, N. (2), & Amram, O. (3). (n.d.). The need for GIScience in mapping COVID-19. *Health and Place.* https://doi-org.libproxy.txstate.edu/10.1016/j.healthplace.2020.102389

Sha, D., Malarvizhi, A.S., Liu, Q., Tian, Y., Zhou, Y., Ruan, S., Dong, R., Carte, K., Lan, H., Wang, Z. and Yang, C., 2020. A State-Level Socioeconomic Data Collection of the United States for COVID-19 Research. Data, 5(4), p.118

Smith, C. D., & Mennis, J. (2020). Incorporating Geographic Information Science and Technology in Response to the COVID-19 Pandemic. *Preventing Chronic Disease, 17, E58.*

https://doi-org.libproxy.txstate.edu/10.5888/pcd17.200246

S.M. Moghadas, A. Shoukat, M.C. Fitzpatrick, C.R. Wells, P. Sah, A. Pandey, *et al.* Projecting hospital utilization during the COVID-19 outbreaks in the United States. *Proc. Natl. Acad. Sci.,* 117 (2020), pp. 9122-9126, 10.1073/pnas.2004064117

Tobler, W. R., 1970. A computer movie simulating urban growth in the Detroit region. Econ. Geogr. 46, 234-240.

Yang, C. ( 1,2 ), D. ( 1,2 ) Sha, Q. ( 1,2 ) Liu, Y. ( 1,2 ) Li, H. ( 1,3 ) Lan, Z. ( 1,6,7 ) Zhang, Z. ( 1,2 ) Wang, et al. 2021. "Taking the Pulse of COVID-19: A Spatiotemporal Perspective." *International Journal of Digital Earth* 13 (10): 1186–1211. Accessed January 27. doi:10.1080/17538947.2020.1809723.

Worldometer. The United State Coronavirus2020. https://www.worldometers.info/coronavirus/country/us/ (accessed February 28, 2021).

Wu, D. (2020). Spatially and temporally varying relationships between ecological footprint and influencing. factors in China's provinces Using Geographically Weighted Regression (GWR). *Journal of Cleaner Production*, *261*. https://doi-org.libproxy.txstate.edu/10.1016/j.jclepro.2020.121089

"BC THE ED SHOW with ED S 01." 2011, July 14. https://search- ebscohost.coom/login.aspx?direct=true&db = apg&AN=dee9d391ee174b16bd70d0e720fd23e6&site=eds-live&scope=site.

Himmelstein DU, Woolhandler S, Cooney R, McKee M, Horton R. The Lancet Commission on public policy and health in the Trump era. *Lancet.* 2018 Sep 22;392(10152):993-995. doi: 10.1016/S0140-6736(18)32171-8. Epub 2018 Sep 20. PMID: 30264712.

Huajie Jin, Haiyin Wang, Xiao Li, et al. Economic burden of COVID-19, China, January-March, 2020: a cost-of-illness study. *Bulletin of the World Health Organization*. 2021;99(2):112-124. doi:10.2471/BLT.20.267112

Zheming Yuan, Yi Xiao, Zhijun Dai, Jianjun Huang, Zhenhai Zhang, and Yuan Chen. 2020. "Modelling the Effects of Wuhan's Lockdown during COVID-19, China." *Bulletin of the World Health Organization* 98 (7): 484–94. doi:10.2471/BLT.20.254045.

Megaloikonomos, P.D. ( 1 ), M. ( 2 ) Thaler, I. ( 2 ) Khosravi, V.G. ( 3 ) Igoumenou, T. ( 4 ) Bonanzinga, M. ( 5 ) Ostojic, A.F. ( 6 ) Couto, and J. ( 7 ) Diallo. 2021. "Impact of the COVID-19 Pandemic on Orthopaedic and Trauma Surgery Training in Europe." *International Orthopaedics* 44 (9): 1611–19. Accessed February 6. doi:10.1007/s00264-020-04742-3.

Menut, L. ( 1 ), B. ( 1,2 ) Bessagnet, S. ( 1 ) Mailler, R. ( 1 ) Pennel, A. ( 1 ) Cholakian, and G. ( 3 ) Siour. 2021. "Impact of Lockdown Measures to Combat Covid-19 on Air Quality over Western Europe." *Science of the Total Environment* 741. Accessed February 6. doi:10.1016/j.scitotenv.2020.140426.

Blue, Sarah A.; Devine, Jennifer A.; Ruiz, Matthew P.; McDaniel, Kathryn; Hartsell, Alisa R.; Pierce, Christopher J.; Johnson, Makayla; Tinglov, Allison K.; Yang, Mei; Wu, Xiu; Moya, Sara; Cross, Elle; Starnes, Carol A. 2021. "Im/Mobility at the US–Mexico Border during the COVID-19 Pandemic" *Soc. Sci.* 10, no. 2: 47. https://doi.org/10.3390/socsci10020047

Lee D, Choi B. Policies and innovations to battle Covid-19 – A case study of South Korea. *Health Policy and Technology*. 2020;9(4):587-597. doi:10.1016/j.hlpt.2020.08.010

Sandeep Grover, Aseem Mehra, Swapnajeet Sahoo, Ajit Avasthi, Adarsh Tripathi, Avinash D'souza,

Gautam Saha, et al. 2020. “State of Mental Health Services in Various Training Centers in India during the Lockdown and COVID-19 Pandemic.” ***Indian Journal of Psychiatry*** 62 (4): 363–69. https://search-ebscohost-com.libproxy.txstate.edu/login.aspx?direct=true&db=lah&AN=20203452166&site=eds-live&scope=site.

Mellish Timothy I., Luzmore Natalie J., Shahbaz Ahmed Ashfaque. Why were the UK and USA unprepared for the COVID-19 pandemic? The systemic weaknesses of neoliberalism: a comparison between the UK, USA, Germany, and South Korea. ***Journal of Global Faultlines***. 2020;7(1):9-45. doi:10.13169/jglobfaul.7.1.0009

Ros, Francisco, Rebecca Kush, Charles Friedman, Esther Gil Zorzo, Pablo Rivero Corte, Joshua C. Rubin, Borja Sanchez, Paolo Stocco, and Douglas Van Houweling. 2021. “Addressing the Covid-19 Pandemic and Future Public Health Challenges through Global Collaboration and a Data-driven Systems Approach.” ***Learning Health Systems*** 5 (1): 1–12. doi:10.1002/lrh2.10253.

Guliyev, Hasraddin. 2020. “Determining the Spatial Effects of COVID-19 Using the Spatial Panel Data Model.” ***Spatial Statistics*** 38 (August). doi:10.1016/j.spasta.2020.100443.

Luo Y, Yan J, McClure S. Distribution of the environmental and socioeconomic risk factors on COVID-19 death rate across continental USA: a spatial nonlinear analysis. ***Environmental Science and Pollution Research***. 2021;28(6):6587. doi:10.1007/s11356-020-10962-2

McNeil, L. M., & Kelso, T. S. (2013). ***Spatial temporal information systems : an ontological approach using STK®***. CRC Press.

Zou H, Xue L. A Selective Overview of Sparse Principal Component Analysis. ***Proceedings of the IEEE, Proc IEEE***. 2018;106(8):1311-1320. doi:10.1109/JPROC.2018.2846588

Gadicherla S, Krishnappa L, Madhuri B, et al. Envisioning a learning surveillance system for tuberculosis. ***PLoS ONE***. 2020;15(12):1-14. doi:10.1371/journal.pone.0243610

https://www.dallasnews.com/news/2021/01/22

# Figures

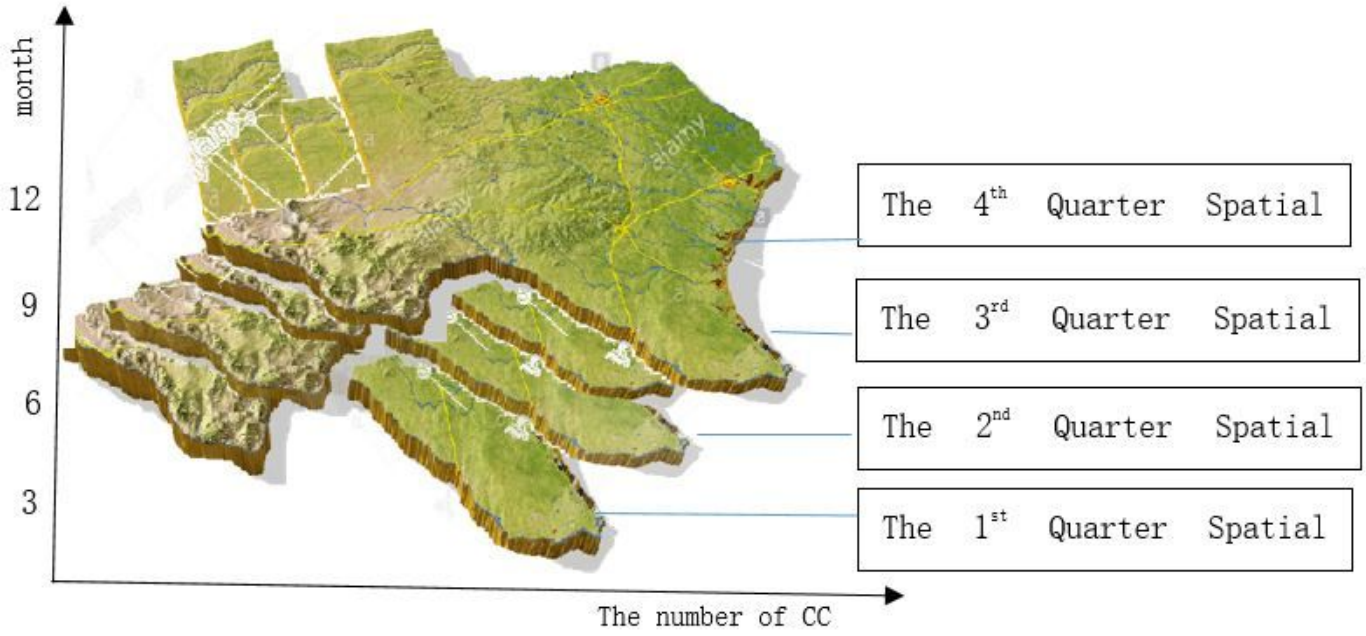



**Figure 1**

Temporal-Study Framework

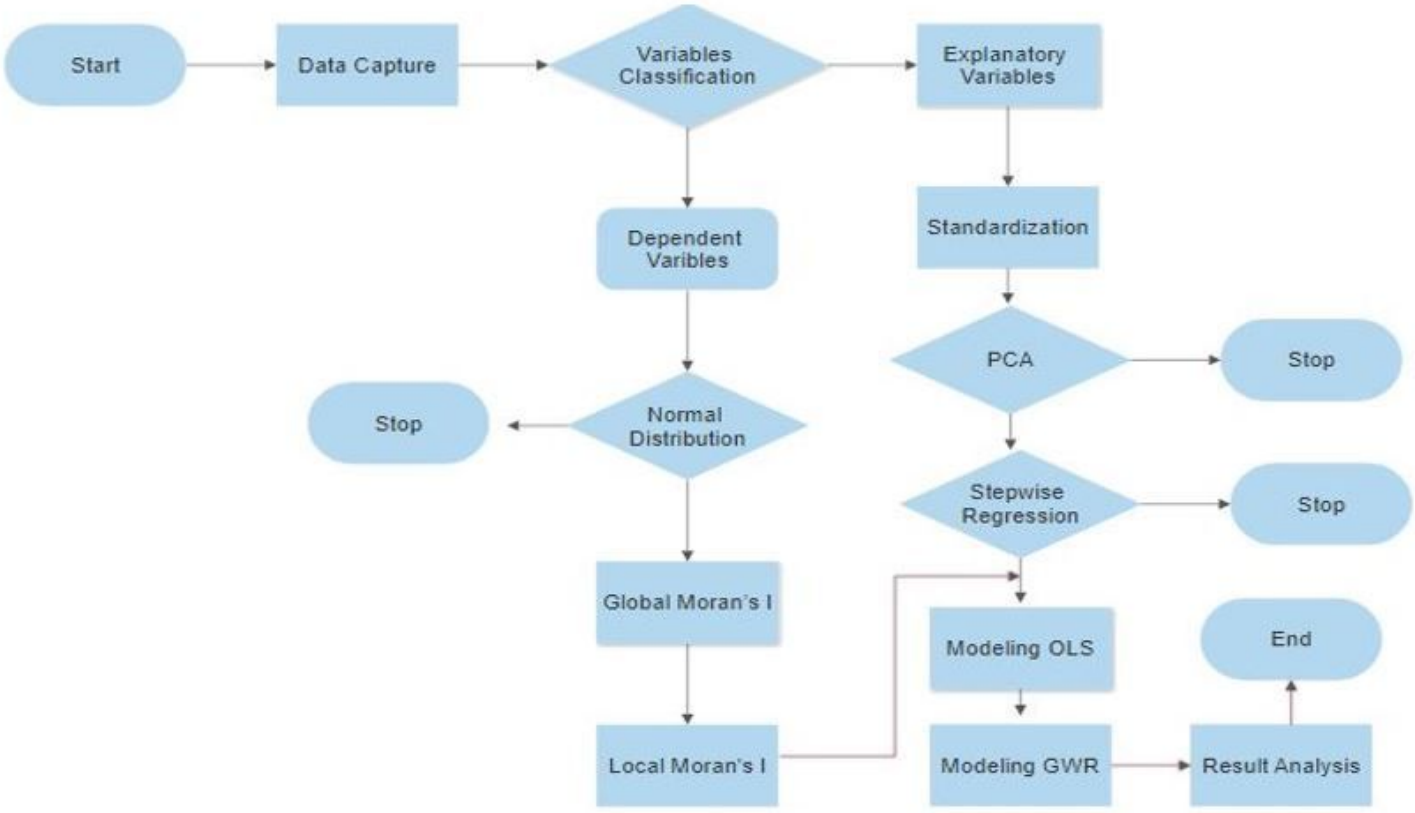



**Figure 2**

Data Flow

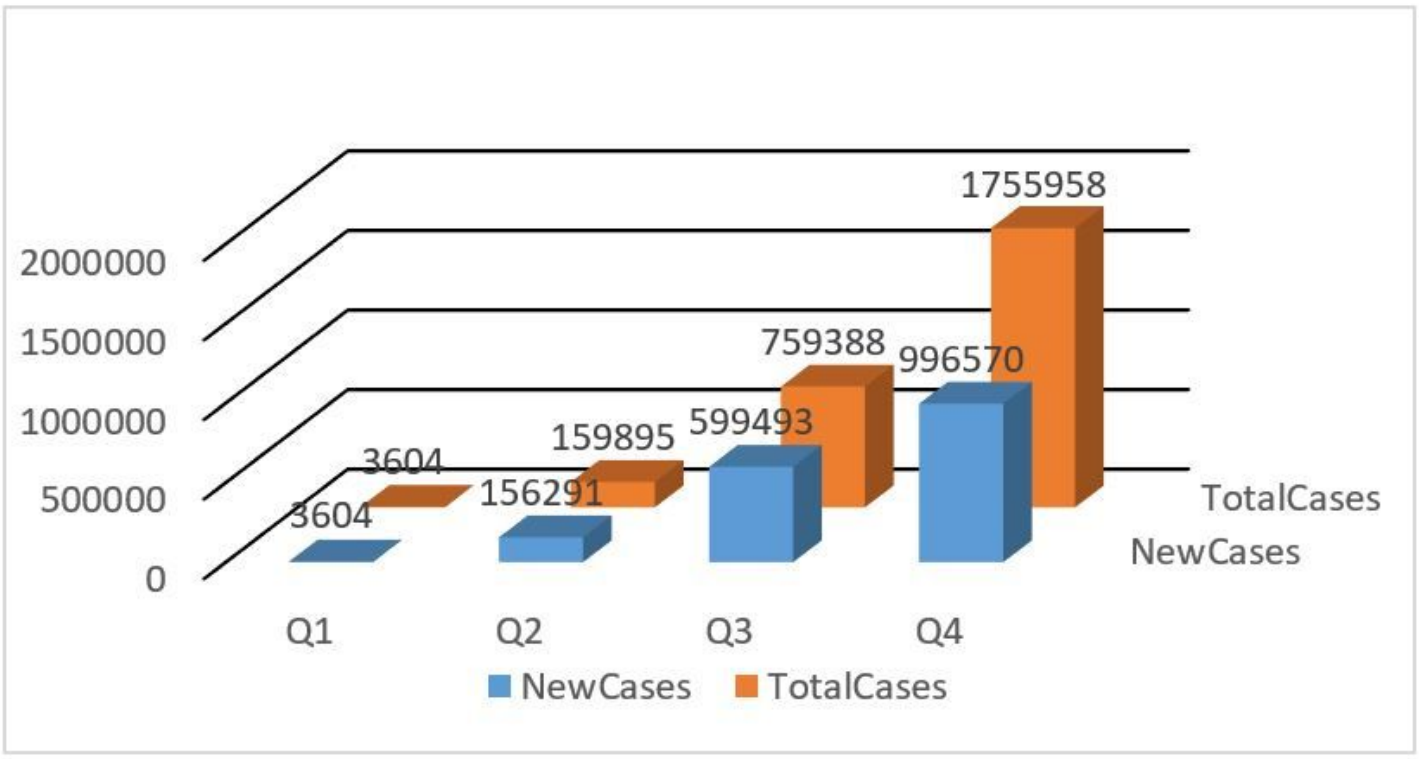



**Figure 3**

Texas Cases Changes over time in 2020

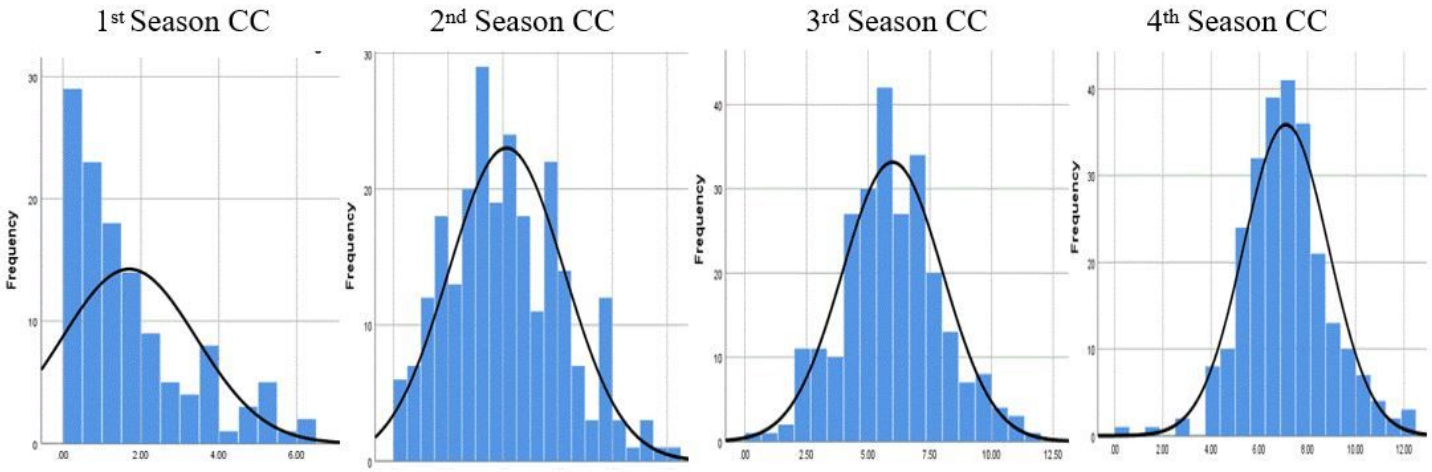



**Figure 4**

CC distribution graph

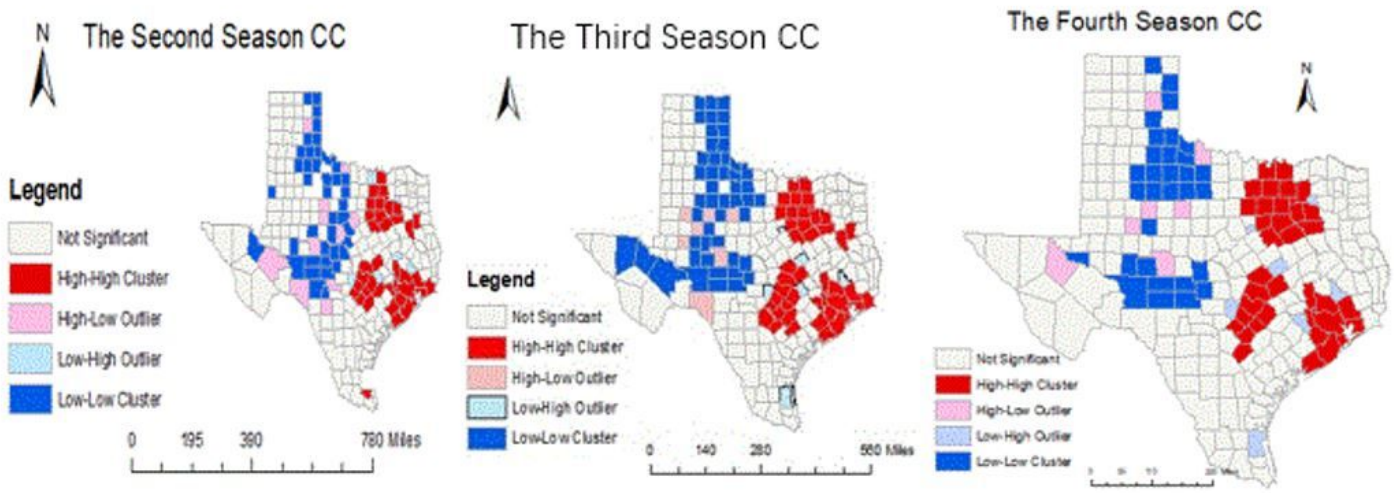
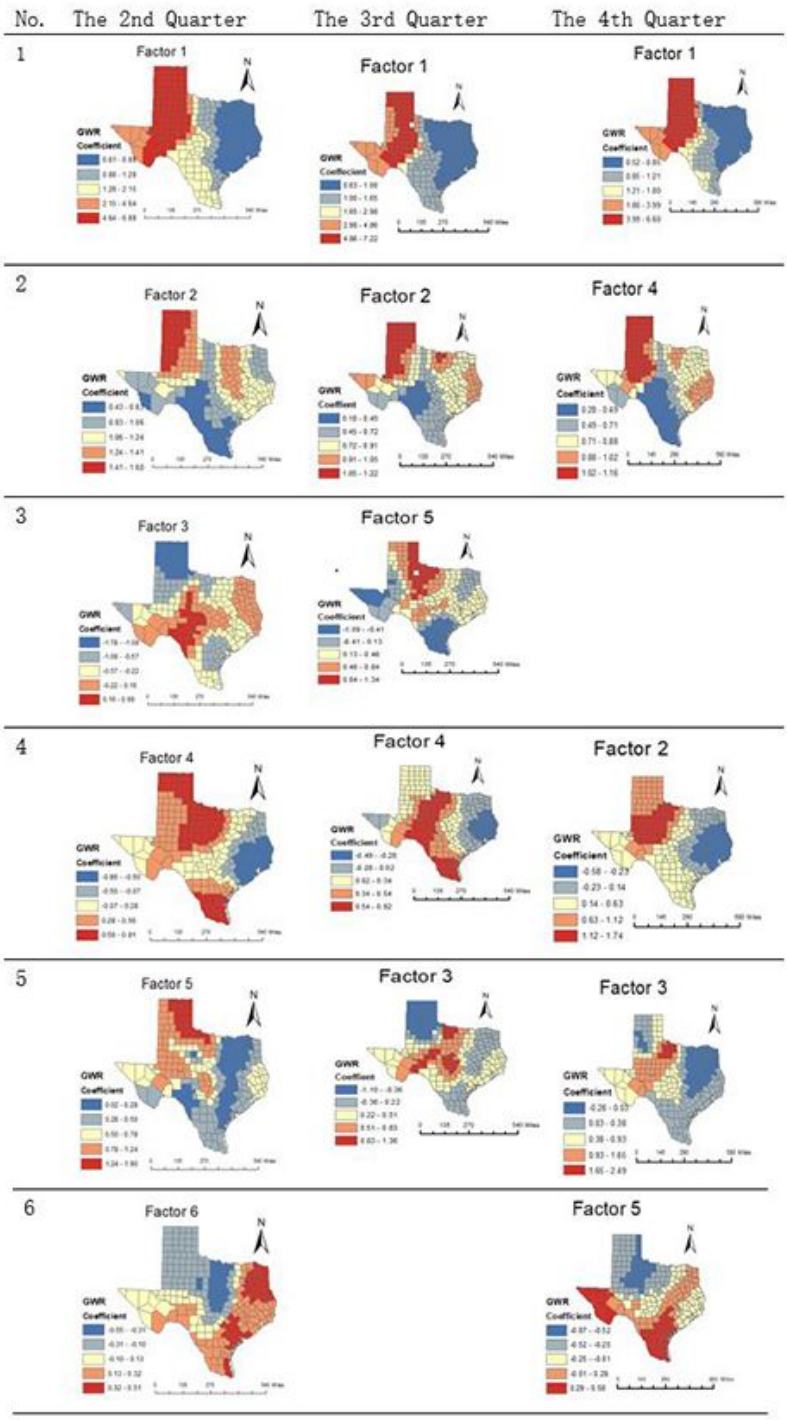



**Figure 5**

Local Moran's Model of CC




## Figure 6

Factors spatial-temporal distribution of CC in the GWR Model