



# Exploration of spatial-temporal varying impacts on COVID-19 cumulative case in Texas using geographically weighted regression (GWR)

Xiu Wu<sup>1</sup> · Jinting Zhang<sup>2</sup>

Received: 18 December 2020 / Accepted: 22 March 2021 / Published online: 10 April 2021

© The Author(s), under exclusive licence to Springer-Verlag GmbH Germany, part of Springer Nature 2021

## Abstract

Since COVID-19 is extremely threatening to human health, it is significant to determine its impact factors to curb the virus spread. To tackle the complexity of COVID-19 expansion on a spatial–temporal scale, this research appropriately analyzed the spatial–temporal heterogeneity at the county-level in Texas. First, the impact factors 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 uses COVID-19 cumulative case as the dependent variable and the common factors as the independent variables. According to the virus prevalence hierarchy, the spatial–temporal disparity is categorized into four quarters in the GWR analysis model. The findings exhibited that GWR models provide higher fitness and more geodata-oriented information than OLS models. In El Paso, Odessa, Midland, Randall, and Potter County areas in Texas, population, hospitalization, and age structures are presented as static, positive influences on COVID-19 cumulative cases, indicating that they should adopt stringent strategies in curbing COVID-19. Winter is the most sensitive season for the virus spread, implying that the last quarter should be paid more attention to preventing the virus and taking precautions. This research is expected to provide references for the prevention and control of COVID-19 and related infectious diseases and evidence for disease surveillance and response systems to facilitate the appropriate uptake and reuse of geographical data.

**Keywords** Geographically weighted regression (GWR) · Spatial-temporal varying impacts · COVID-19 cumulative case

## 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, 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 people’s daily lives and reshapes people’s routine behaviors, especially a pervasive sense of quarantine fatigue and panic attacks of getting infected are challenging human fortitudes (Ahmar and Boj 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 (Huajie et al. 2021; Yuan et al. 2020; P.D. et al., 2020; Menut et al. 2021). The USA was called the first country with the highest confirmed cases of COVID-19 in the world (Ahmar and Boj 2020; Worldometer. The United States coronavirus 2020). Until February 5, 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 USA. A severe economic downturn behind figures was predicted by considering how policy has supported businesses and families

---

Responsible Editor: Philippe Garrigues

✉ Jinting Zhang  
whuzjt@whu.edu.cn

Xiu Wu  
x\_w10@txstate.edu

<sup>1</sup> Department of Geography, Texas State University, 601 University drive, San Marcos 78666, TX, USA

<sup>2</sup> School of Resource and Environmental Science, Wuhan University, 129 Luoyu Rd., Wuhan 430079, Hubei, China

since March 2020. The first COVID-19 case in the USA was confirmed on January 19, 2020, in Washington State (Harcourt et al. 2020; Holshue et al., 2020; Ellis 2020; Qu, et al., 2020). From summer to winter in 2020, the virus, as a perfect storm, virtually spreads on every part of the USA at a speed unprecedented in American history, according to Johns Hopkins University data. Ducharme 2021 argued the pandemic had claimed more than three times the American lives that were lost in the Vietnam War. The coast-to-coast surge is causing hospitals across the country to the edge of catastrophe. Doctors and nurses are exhausted, and their intensive-care units are 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 the area except for Alaska and California. The state extends nearly 1000 miles (1600 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 USA, Texas plays a significant role in curbing the spread of COVID-19 in the USA. Therefore, it is plausible to choose Texas as our study object of COVID-19. Ostensibly, the Texas government made stringent policy interventions to mitigate the spread of COVID-19 based on unremitting Texas Executive Orders (TEO) and Public Health Disaster Declarations (PHDD). On March 16, 2020, the USA promulgated the President's Coronavirus Guidelines for America, calling upon Americans to slow down 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 4, the second phase on May 18, the third phase on June 3, the fourth phase on July 4, and the fifth phase on December 15 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 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 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 from March 19, 2020, to January 15, 2021. "The novel coronavirus (COVID-19) has been recognized globally as a contagious respiratory virus" mentioned on the 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, and from elementary schools to universities (Sha et al. 2000). The socioeconomic impact of COVID-19 is well documented as well (Bashir et al. 2020a, 2020b; Sha et al. 2020a, 2020b; Yang et al. 2021). When tracing the COVID-19 footprint, the first case was announced by the Texas Department of State Health Services on March 4 in Fort Bend County. Texas surpasses 2,433,110 total COVID-19 cases and 128,000 death 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 thought that not controlling the virus spread is mainly responsible for a potential policy from the Trump Administration, which has done little to counter the rampant misinformation about the pandemic and has made numerous incorrect statements about the virus's origins, spread, and deadliness. Indeed, it is worth noting that establishing a long-lasting, real-time, and dynamic emergency alert system on 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 areas of vulnerability, and most vulnerable groups based on the county level. The objective of this paper is to investigate the spatial-temporal variability between population (age structure, race, and gender) and COVID-19 cumulative cases at 254 Texas counties, in the context of considering the impacts of social-economic (unemployment and annual income) and environmental factors, via spatial stratified heterogeneity analysis using geographically weighted regression (GWR) models. This research will facilitate the 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 (Cámara 2020).

Under the 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 Blue et al. 2021 regarding asylum seekers' immobility/mobility at the US-Mexico Border during the COVID-19 pandemic. There are 482,958 academic journals that published COVID-19 topics. In light of Web of Science, 567 articles are being published in 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 (Lee and Choi 2020). COVID-19 pandemic and lockdown in India have led to the collapse of regular mental health services (Grover et al. 2020). Interestingly, a comparison about COVID-19 response between the UK, USA, Germany, and South Korea is rooted in the systemic weaknesses of neoliberalism (Megaloiconomos et al. 2021; Mellish et al. 2020). Mellish et al. (2020) proposed that economic recession and austerity impeded health care investment in the UK and USA. Air quality impacts of COVID-19 are mentioned in Europe and China (Mellish et al. 2020; Menut et al. 2021; 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 (Ros et al. 2021).

Since the COVID-19 spreading represented geographical dependence, geographic information systems (GIS) can combine divergent spatial datasets 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 and Mennis 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 geographical inequity and a long-term impact on global health accurately, regardless of linear or nonlinear regressions (Cássaro and Pires 2020; Rosenkrantz et al. 2020; Smith and Mennis 2020; Guliyev 2020). For instance, Ahmar and Boj 2020 predicted COVID-19 confirmed cases in the USA with the SutteARIMA method. Luo et al. (2020) exhibited a spatial nonlinear analysis of the COVID-19 death rate in the USA. Unfortunately, GIS spatial-temporal analysis is seldom mentioned. Chaowei Yang et al. (2021) 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. 2021). 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 et al. 2020; Smith and Mennis 2020; Lakhani 2020; Rosenkrantz et al. 2020). Different methods are used to observe the goodness-of-fit test of the regression (e.g., multiple geographically 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 USA (Bag et al. (2020); Carteni et al. 2021; 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. 2020a, 2020b), longitudinal detection and explanation of COVID-19 in the USA are essential for dealing with the linear scenario in a local regression model. On the other hand, according to an 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 factors such as race/ethnicity and socioeconomic status, leading to encode the vulnerability to adverse health outcomes such as negative effects of COVID-19.

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

## Methodology

### Data source

Demographical factors such as age structure, gender, and race are examined to detect the most vulnerable groups. Since PM<sub>2.5</sub> and Air Quality Index (AQI) are used to measure the severity of air pollution, which is highly related to respiratory diseases, PM<sub>2.5</sub> and AQI are also considered explanatory variables. COVID-19 data (cumulative cases and new cases) as dependent variables are from the Centers for Disease Control and Prevention (CDC), COVID-19 fatality data based on death certificates. A fatality is counted as a COVID-19 fatality when the medical certifier attests to the death certificate that COVID-19 is a cause of death. Mortality is equal to fatalities divided by cumulative cases. Hospitalization (i.e., THB and BPC) from the Texas Department of State Health Services (DSHS) is reported daily by hospitals through eight

Hospital Preparedness Program providers that coordinate health care system preparedness and response activities in Texas. They are viewed as explanatory variables over 2020. Population data (e.g., race, diverse age group, gender, and population density) as explanatory variables are from the 2019 US Census Bureau. Economic data (e.g., annual income) as explanatory variables are from the Texas Association of Counties, and the statistical period is 2019. Environmental data is the result through interpolating limited samples, which are from the US Environmental Protection Agency (i.e., AQI and PM2.5) and National Weather Service (i.e., temperature and precipitation); statistical period is the whole year of 2020, categorizing four quarters. Quarterly data are real-time raw data at the end day of each quarter (Table 1).

**Study area**

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

**Spatial–temporal study framework**

For the temporal study in the paper, time series models were classified into four layers according to four seasons in 2020. Quarterly statistical data are based on environmental and 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.

From a 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 least squares (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.

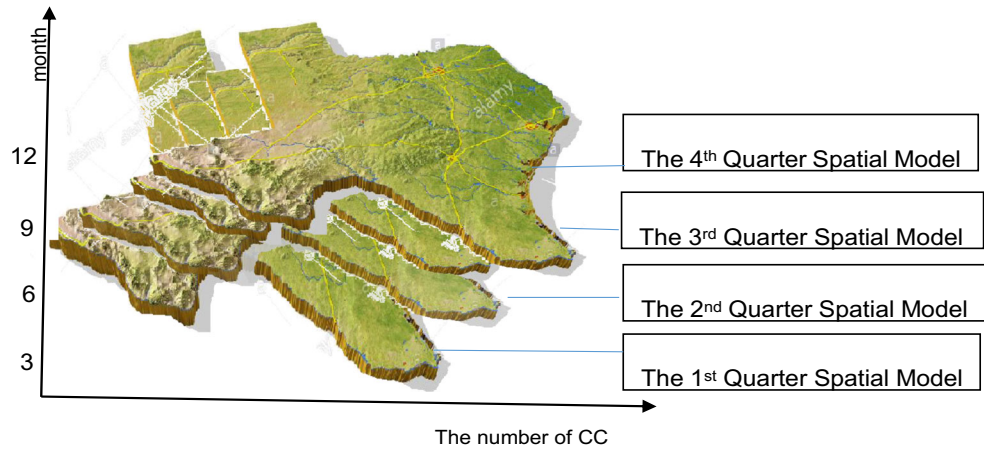
**Data preprocessing and preparing**

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 a surge without turn points. That means the

**Table 1** A list of variables used for geostatistical analysis

Variable category	Variable name	Acronym	Variable description
Economic	Annual income	PCI	Annual income per 1000 residents
	Unemployment	UEM	Percent of residents who do not 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
Demographic	Land area	LA	Total land area per county
	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
	Population between 20 and 59	P59	Percent of residents who are between 20 and 59
Health	Population beyond 80	P80	Percent of residents who are beyond 80
	Total hospital beds	THB	Total hospital beds
	Beds per capital	BPC	Incidents per 1000 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

**Fig. 1** Temporal study framework



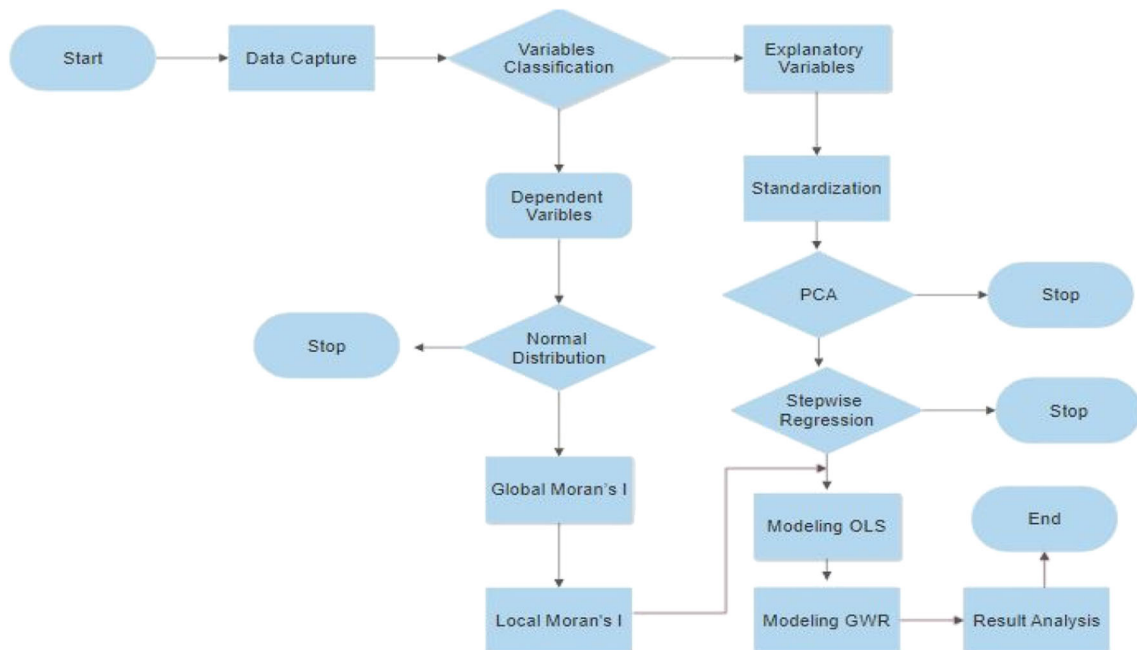
spread of COVID-19 is monotonically increasing without controlling.

Data standardization is the process of making sure that your dataset can be compared to other datasets. 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

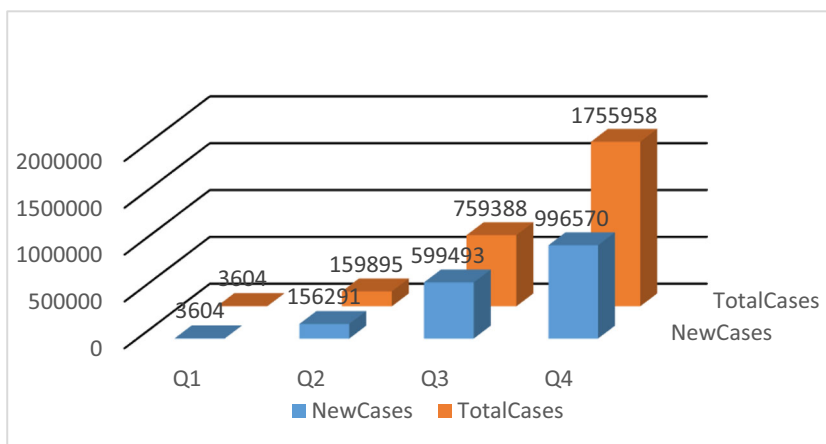
minimize information loss. It does so by creating new uncorrelated variables that successively maximize variance. In the PCA procedure, a set of possibly correlated variables is transformed into a set of linearly uncorrelated variables using the orthogonal transformation. This set of linearly uncorrelated 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 (Gray 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



**Fig. 2** Data flow

**Fig. 3** Texas cases changes over time in 2020



variables improve the model to a statistically significant extent (Guidolin and Pedio 2020).

## Method

### Ordinary least squares

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,  $Y_i$  and  $X_i$ , of the form

$$Y_i = \alpha + \beta X_i + u_i \quad i = 1, 2, \dots, n \tag{1}$$

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

### Geographically 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 (GWR) model (Fotheringham and Charlton 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).

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

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

$$(u_i, v_i) = (X^T W(u_i, v_i) X)^{-1} X^T W(u_i, v_i) y \tag{3}$$

where  $W$  is the spatial weight matrix, whose selection and setting are the core issues 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 Gaussian. 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

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

where  $w_{ij}$  represents the distance weight from sample  $i$  to sample  $j$ ;  $d_{ij}$  is the Euclidean distance between sample  $i$  and sample  $j$ ;  $\theta$  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). Since it is more suitable for the Gaussian model and its robust performance, the CV method was adopted in the article to determine the optimal bandwidth. At the same time, the ANOVA test method was used to test the null hypothesis that the GWR model has no improvement over the OLS model. Finally, the statistical significance of all local regression coefficients was determined using the pseudo  $t$  values calculated by GWR (Clement et al. 2009).

## Results and findings

### 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. The first is that the uncertain variable is symmetric about the mean and the other is that the 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 as a normal distribution except for the first quarter in Fig. 4. When modeling GWR regressions, the first quarter CC is overlooked as a skewed distribution.

### 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, and P59 and negatively significant to P80. The third quarter CC is

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

### 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 into four clusters including HH, HL, LH, and LL. HH is distributed in northern Texas at 39 counties, HL is distributed at 11 counties, LH is limited at 6 counties, and LL occupied the south and east of Texas at 41 counties. In the third season, HH is distributed in northern Texas at 44 counties, HL is distributed at 7 counties, LH is limited at 6 counties, and LL is distributed in the south and east of Texas at 48 counties. In the fourth season, HH is distributed in northern Texas at 37 counties, HL is distributed at 7 counties, LH is limited at 7 counties, and LL is distributed in the south and east of Texas at 31 counties.

### 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 ( $p < 0.001$ ),

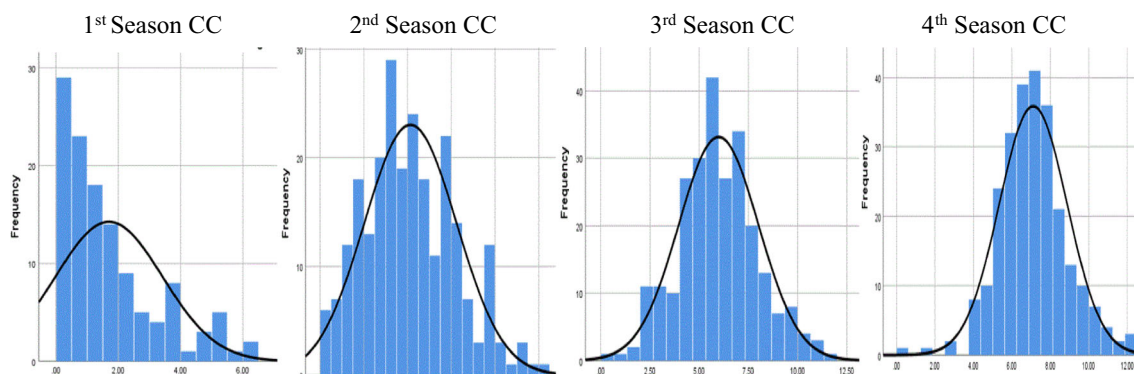


Fig. 4 CC distribution graph Himmelstein et al. 2020

**Table 2** Person correlation between CC and explanatory variable

Explanatory variables	CC quarter 1	CC quarter 2	CC quarter 3	CC quarter4
	Coef./Sig.	Coef./Sig.	Coef./Sig.	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

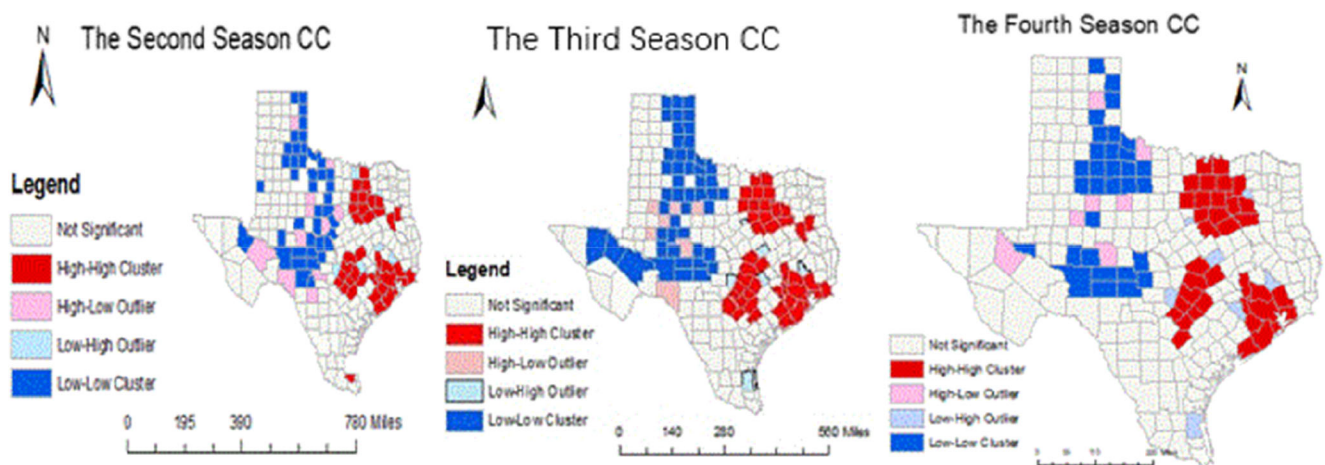
Note: \*Correlation is significant at the 0.05 level (2 tailed). \*\*Correlation is significant at the 0.01 level (2 tailed).

validating that the correlation between variables does exist in the GWR models.

Communality is a common variance between 0 and 1, using the remaining variables as factors, and was used to determine if any variables should be excluded from the factor analysis. A 0.7 threshold is used to determine the significance of explanatory variables (David et al. 2012).

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 and 5 components in the third quarter and the fourth quarter are produced, explaining cumulative 76.57% (the second quarter), 70.60% (the third quarter), and 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 is positively related to the 20–59 population but negatively related to the 80 population. That means the 20–59 population directly contributes to COVID-19 patients, and 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 plays a positive role 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



**Fig. 5** Local Moran’s model of CC



risk of COVID-19 infection) and 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 keeping spatial distancing helps in COVID-19 CC reduction. Factor 6 refers to medical supplies (i.e., BPC), meaning hospital beds are positively related to COVID-19 CC. Five factors in the third quarter are identical with the 5 factors in the first quarter except for factor 6. Five factors in the last quarter are similar to 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 the increase of COVID-19. The concrete relationships are shown in Tables 3 and 4.

### Comparison of composite OLS and GWR models

Modeling OLS is used to examine whether there is a linear relationship between CC and its factors. By the *t* test and *f* test, all factors have passed. Modeling GWR is used 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, the 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 has decreased from 883.74 in the OLS model to 811.99 in the GWR, and  $R^2$  has changed from 0.54 in the OLS model to 0.77 in the GWR model. Neighbors are declined from 254 neighbors in the OLS model to 77 neighbors in the GWR model. In the third quarter, the AICc value has decreased from 870.29 in the OLS model to 790.31 in the GWR, and  $R^2$  has changed from 0.55 in the OLS model to 0.77 in the GWR model. Neighbors are declined from 254 neighbors in the OLS model to 77 neighbors in the GWR model. In the fourth quarter, the AICc value has decreased from 850.42 in the OLS model to 778.75 in the GWR, and  $R^2$  has 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 (Table 5).

### GWR result analysis

#### Spatial change of CC factors

Based on existing research, COVID-19 quarterly GWR models are also implemented in the research area (Liu et al. 2020; Mollalo et al. 2020). Figure 6 incorporates Texas spatial–temporal distribution maps based on 6 factors in terms of 6 aspects in Table 3 in three quarters.

In the second quarter, factor 1 among the 6 factors has the largest effects on CC in northern Texas, thanks to the maximum coefficient of 6.88. It has the lowest impact in eastern Texas due to the coefficient range of 0.61–0.88. indicating total population and hospitalization are the key factors of COVID-19 and northern Texas is the main precaution and control area of COVID-19. Factor 2 (age structure) positively affects COVID-19 spatial heterogeneity in central Texas with pink color. The area of the largest coefficient range 1.41–1.6 is distributed in northern Texas. The smallest impacts of the coefficient range 0.43–0.83 are the coastal area at the bottom of the map. Factor 3 is the air quality index, having remarkable spatial disparity for its coefficient from a range of –1.78–1.09 to a range of 0.15–0.66. In central Texas, the improvement of air quality is driven by COVID-19 CC, but it reversely works in northern Texas. That indicates that AQI has a spatial non-stationary and environmental harness is available, reducing CC in northern Texas. Factor 4 is an economic composite index with a coefficient range from –0.86–0.50 to range 0.56–0.81. The spatial heterogeneity is located among northern Texas, coastal counties, and eastern Texas. Factor 5 is the natural supply index with a coefficient range from 0.02–0.28 to range 1.24–1.90. The spatial heterogeneity is subtle. Factor 6 is the medical supply index with a coefficient range from –0.56–0.31 to range 0.32–0.51. It is evident to see that the change of spatial heterogeneity and the medical condition in

**Table 3** Factor roles in modeling OLS and GWR regressions

No.	Items	Quarter 2	Quarter 3	Quarter 4
1	Population and hospitalization	Factor 1 (THB, POD, TP)	Factor 1 (THB, POD, TP)	Factor 1 (THB, POD, TP)
2	Age structure	Factor 2 (P59, P80)	Factor 2 (P59, P80)	Factor 4 (P59, P80)
3	Air quality	Factor 3 (TPE, AQI)	Factor 5 (AQI)	
4	Economic	Factor 4 (PCI, UEM)	Factor 4 (PCI, UEM)	Factor 2 (PCI, UEM)
5	Natural supply	Factor 5 (LA)	Factor 3 (LA)	Factor 3 (PCN, LA)
6	Medical supply	Factor 6 (BPC)		Factor 5 (BPC)

**Table 4** Rotated component matrix

Variables	The second quarter component						The third quarter component					The fourth quarter component							
	2		3		4		5		6		Extract	1	2	3	4	5			
	1	2	1	2	1	2	1	2	1	2	3	4	5	1	2	3	4	5	
TPE	0.771	0.140	-0.001	<b>0.826</b>	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	<b>0.769</b>	0.056	-0.163
AQI	0.742	-0.023	-0.023	<b>0.828</b>	-0.006	-0.181	-0.149	0.633	0.294	0.021	-0.191	0.031	<b>0.713</b>	0.504	0.175	0.551	0.24	-0.022	-0.334
THB	0.956	<b>0.973</b>	0.083	0.021	-0.022	0.003	0.043	0.952	<b>0.974</b>	0.056	0.013	-0.015	-0.006	0.958	<b>0.974</b>	0.012	0.024	0.082	0.044
BPC	0.779	0.108	-0.099	-0.090	-0.010	0.070	<b>0.863</b>	0.343	0.148	0.04	0.048	0.081	-0.558	0.642	0.121	-0.003	0.085	-0.053	<b>0.786</b>
POD	0.927	<b>0.942</b>	0.158	0.033	-0.053	0.097	-0.020	0.926	<b>0.945</b>	0.124	0.1	-0.056	0.072	0.926	<b>0.942</b>	-0.013	0.115	0.151	-0.039
LA	0.755	0.083	0.072	-0.111	0.198	- <b>0.831</b>	-0.018	0.71	0.069	0.062	- <b>0.804</b>	0.18	0.148	0.613	0.077	0.195	- <b>0.747</b>	0.1	0.034
PCI	0.658	0.145	0.076	-0.015	- <b>0.782</b>	0.109	-0.083	0.69	0.138	0.046	0.097	- <b>0.807</b>	0.061	0.626	0.177	- <b>0.728</b>	0.093	0.071	-0.227
TP	0.972	<b>0.978</b>	0.117	0.038	-0.029	0.008	-0.012	0.967	<b>0.978</b>	0.08	0.019	0.514	0.262	0.972	<b>0.979</b>	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	<b>0.815</b>	0.129	-0.115	0.676	0.025	0.004	0.126	<b>0.801</b>	0.136	0.658	0	<b>0.806</b>	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	<b>0.851</b>	0.064	-0.094	0.172	0.053	0.786	0.185	<b>0.835</b>	0.196	-0.101	0.075	0.777	0.17	-0.077	0.17	<b>0.844</b>	-0.034
P80	0.700	-0.189	- <b>0.805</b>	0.087	-0.047	0.065	0.041	0.646	-0.209	- <b>0.774</b>	0.045	-0.025	-0.022	0.688	-0.188	-0.034	0.052	- <b>0.805</b>	0.028

northern Texas are worse than in other Texas counties (<https://www.dallasnews.com/news/2021/01/22>).

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

In the fourth quarter, factor 1 among 6 factors is still the dominant effect on CC without the range of maximum coefficient 3.99–6.6. Spatial heterogeneity is small, implying that it is a fixed factor. Factor 2 is an economic composite index with a coefficient range from -0.58 slight 0.23 to range 1.12 slight 1.74. The spatial heterogeneity is that areas of positive impacts have decreased while areas of negative impacts have increased. Factor 3 is a natural supply index with a coefficient range from -0.28–0.03 to range 1.65–2.49. The spatial heterogeneity is that areas of positive impacts have decreased while areas of negative impacts are moved from northern Texas to eastern Texas. Factor 4 is the age structure index with a coefficient moved from range 0.28–0.49 to range 1.02–1.16. The spatial heterogeneity is that both areas of positive impacts and negative impacts have increased. Factor 5 is the air quality index with a coefficient range from -1.09–0.41 to range 0.84–1.34. Spatial heterogeneity is obviously seen in the change of spatial heterogeneity where positive impacts are from western Texas to eastern Texas while negative impacts are from northern Texas to western and southern Texas. Factor 6 is the medical supply index with a coefficient range from -0.87–0.52 to range 0.29–0.58. It is evident to see the change of spatial heterogeneity where areas of positive impacts are moved from eastern Texas to western and southern Texas while areas of negative effects have decreased and moved.

**Table 5** GWR and 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
$R^2$	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
$R^2$	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

### Temporal change of CC factors

Population and hospitalization impacts on COVID-19 within 3 quarters are 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. More importantly, northern Texas, including El Paso, Odessa, Midland, Lubbock, and Amarillo areas, is the most important area in curbing the 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 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 has increased from  $-1.78$ – $0.66$  to  $-1.09$ – $1.34$ . It demonstrated that the role of the environment 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 and 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 has 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 at the bottom of the map is shrinking until it has disappeared in the last quarter. It reveals that policy control and human self-consciousness are beneficial for mitigating COVID-19 spread.

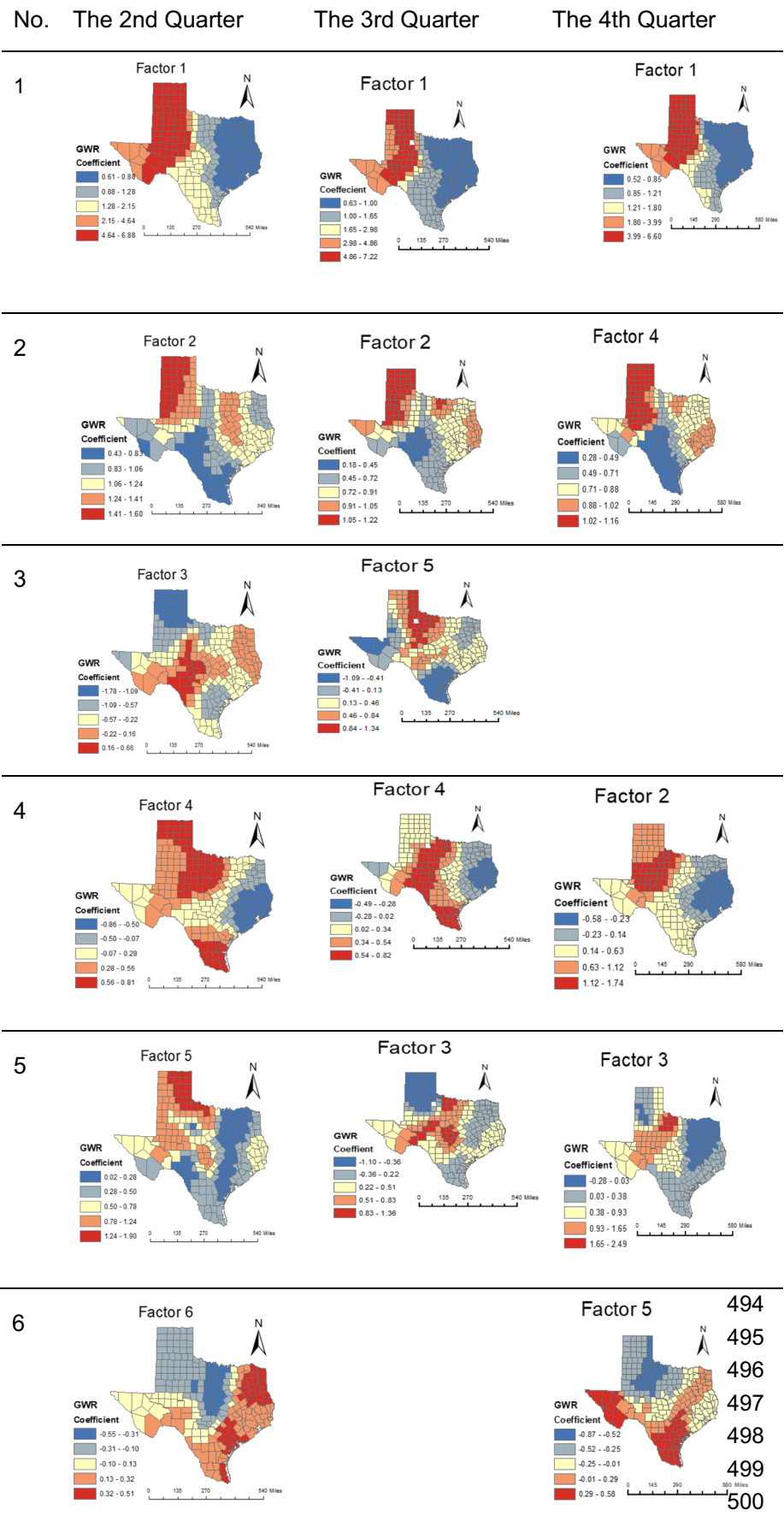
Natural supply impacts in three quarters have fluctuated. First, the coefficient range within three quarters has changed:  $0.02$ – $1.9$  and  $-1.10$ – $1.36$  into  $-0.28$ – $1.49$ . It demonstrated that the role of natural supply is out of control. Second, the cluster of positive impacts with red 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 impact factors.

Medical supply impacts in three quarters have fluctuated as well. First, the coefficient range within two quarters has changed:  $-0.5$ – $0.51$  and  $0$  into  $-0.87$ – $0.58$ . It demonstrated that the role of medical supply impacts is small 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.

### Discussion

In this study, 14 potential risk variables are selected from 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 lacks the consideration of time-series models, spatial–temporal GWR is explored to accurately identify the unbalanced distribution of COVID-19

**Fig. 6** Spatial-temporal distribution factors of CC in the GWR model



cumulative cases and the complex relationship between the COVID-19 CC and its risk factors (Luo et al. 2020). Four quarters in 2020 are categorized to model quarterly GWR 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 have exhibited stable, positive influences on COVID-19 cumulative cases. Climate, natural, economic, and medical conditions have 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 the 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 geographic information systems (GIS) with the use of modeling and simulation skills (McNeil and Kelso 2013).

Previous studies have shown that many social–environmental and economic variables are captured to analyze the distribution of COVID-19 cases and 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 simple 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 and Xue 2018). Especially, composite factors are optimized GWR model 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) The 20–59 population is the main source of cumulative cases with a lower death rate, while over 80 population has lower infection rates and higher COVID death rates. Thus, the over 80 population is the most vulnerable group to COVID-19. (3) Race and gender should be paid no attention in 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.

## Limitations

This research just focuses on the Texas COVID-19 scenario, and the application of the research is limited to other states. There is no chronic disease data to support this research. As explanatory variables, they should be incorporated in future studies, although it is exciting to see diabetes (Gupta et al., 2020) and cardiovascular conditions (Du et al. 2020) as 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 use 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.

## Implications

The COVID-19 pandemic revealed systemic flaws in the food distribution system that failed to protect against hunger, and diet influenced non-communicable diseases. It also exposes the conditions that made people who are living with 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 health-harming 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 (Himmelstein 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 USA 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 the 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.

## Conclusion

Spatial–temporal geostatistical analysis on COVID-19 CC through GWR is selected real-time raw to expose real scenarios in Texas counties. Thus, it is beneficial for the government of Texas and CDC to make appropriate and scientific judgments, target vulnerable communities, allot health care resources, and reshape disease surveillance and response systems, which remain the core of modern public health practices (Bashir et al. 2020a, 2020b; Gadicherla et al. 2020).

**Acknowledgements** This research is obtained by Dr. F. Benjamin Zhan's support at the Literature review part and Dr. Edwin T. Chow's help in the Methodology section. Thanks for their efforts.

**Author contribution** Xiu Wu: conceptualization, writing—original draft, formal analysis, methodology, software, formal analysis, investigation, and visualization. Jinting Zhang: conceptualization, validation, writing—review and editing, and supervision.

**Data availability** The datasets used during the current study are available from the corresponding author on reasonable request.

## Declarations

**Ethical approval** Not applicable.

**Consent to participate** Not applicable.

**Consent for publication** The co-authors consent to the publication of this work.

**Competing interests** The authors declare no competing interests.

## References

- Ahmar AS, Boj E (2020) Will COVID-19 confirmed cases in the USA reach 3 million? A forecasting approach by using SutteARIMA method. *Current Res Behav Sci* 1(November):100002. <https://doi.org/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. *Reg Sci Policy Pract* 12(6):1063–1103. <https://doi.org/10.1111/rsp3.12359>
- Bashir MF, Benjiang MA, Shahzad L (2020a) A brief review of socio-economic and environmental impact of COVID-19. *Air Quality, Atmosphere & Health: An International Journal* 13(12):1403. <https://doi.org/10.1007/s11869-020-00894-8>
- Bashir A, Malik AW, Rahman AU, Iqbal S, Cleary PR, Ikram A (2020b) MedCloud: cloud-based disease surveillance and information management system. *IEEE Access*, Access, IEEE 8:81271–81282. <https://doi.org/10.1109/ACCESS.2020.2990967>
- Bilal FL, Bashir MF, Komal B, Tan D (2020) Role of electronic media in mitigating the psychological impacts of novel coronavirus (COVID-19). *Psychiatry Res* 289:113041. <https://doi.org/10.1016/j.psychres.2020.113041>
- Blue SA, Devine JA, Ruiz MP, McDaniel K, Hartsell AR, Pierce CJ, Johnson M, Tinglov AK, Yang M, Wu X, Moya S, Cross E, Starnes CA (2021) Im/mobility at the US–Mexico border during the COVID-19 pandemic. *Soc Sci.* 10(2):47. <https://doi.org/10.3390/socsci10020047>
- Câmara G (2020) On the semantics of big Earth observation data for land classification. *J Spatial Inf Sci* 2020(20):21–34. <https://doi.org/10.5311/JOSIS.2020.20.645>
- Carteni A, Di Francesco L, Martino M (2021) The role of transport accessibility within the spread of the coronavirus pandemic in Italy. *Saf Sci* 133:104999. <https://doi.org/10.1016/j.ssci.2020.104999>
- Cássaro FAM, Pires LF (2020) Can we predict the occurrence of COVID-19 cases? Considerations using a simple model of growth. *Sci Total Environ* 728. <https://doi.org/10.1016/j.scitotenv.2020.138834>
- Clement F, Orange D, Williams M, Mulley C, Epprecht M (2009) Drivers of afforestation in Northern Vietnam: assessing local variations using geographically weighted regression. *Appl Geogr* 29(4):561–576. <https://doi.org/10.1016/j.scitotenv.2020.138834>
- David T. Hickman & T. Edwin Chow (2012) Development of a Composite Model of Quality of Life: A Case Study in Austin, Texas, *GIScience & Remote Sensing* 49(6):802–821. <https://doi.org/10.2747/1548-1603.49.6.802>
- Du H, Wang DW, Chen C (2020) The potential effects of DPP-4 inhibitors on cardiovascular system in COVID-19 patients. *J Cell Mol Med* 24(18):10274–10278. <https://doi.org/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](https://search.ebscohost.com/login.aspx?direct=true&db=a9h&AN=147960785&site=eds-live&scope=site).
- Ellis RE (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 AS, Charlton ME (2002) Geographically weighted regression: the analysis of spatially varying relationships. Wiley, New York, Chris Brunsdon
- Gadicherla S, Krishnappa L, Madhuri B, Mitra SG, Ramaprasad A, Seevan R, Sreeranga SD, Thodika NK, Mathew S, Suresh V (2020) Envisioning a learning surveillance system for tuberculosis. *PLoS One* 15(12):1–14. <https://doi.org/10.1371/journal.pone.0243610> <https://www.dallasnews.com/news/2021/01/22>
- Gray V (2017) Principal component analysis: methods, applications, and technology. Nova Science Publishers, Inc
- Grover S, Mehra A, Sahoo S, Avasthi A, Tripathi A, D'souza A, Saha G et al (2020) State of mental health services in various training centers in India during the lockdown and COVID-19 pandemic. *Indian J Psychiatry* 62(4):363–369. [https://doi.org/10.4103/psychiatry.IndianJPsychiatry\\_567\\_20](https://doi.org/10.4103/psychiatry.IndianJPsychiatry_567_20)
- Guidolin M, Pedio M (2020) Forecasting commodity futures returns with stepwise regressions: Do commodity-specific factors help? *Ann Oper Res*. <https://doi.org/10.1007/s10479-020-03515-w>
- Guliyev H (2020) Determining the spatial effects of COVID-19 using the spatial panel data model. *Spatial Statistics* 38(August):100443. <https://doi.org/10.1016/j.spasta.2020.100443>
- Harcourt J, Tamin A, Lu X, Kamili S, Sakthivel SK, Murray J et al (2020) Isolation and characterization of SARS-CoV-2 from the first US COVID-19 patient *BioRxiv*. <https://doi.org/10.1101/2020.03.02.972935>

- Himmelstein DU, Woolhandler S, Cooney R, McKee M, Horton R (2020) The Lancet Commission on public policy and health in the Trump era. *Lancet* 392(10152):993–995. [https://doi.org/10.1016/S0140-6736\(18\)32171-8](https://doi.org/10.1016/S0140-6736(18)32171-8)
- Jin H, Wang H, Li X, Zheng W, Ye S, Zhang S, Zhou J, Pennington M (2021) Economic burden of COVID-19, China, January–March, 2020: a cost-of-illness study. *Bull World Health Organ* 99(2):112–124. <https://doi.org/10.2471/BLT.20.267112>
- 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. *J Pain Symptom Manag* 60(1):e41–e44. <https://doi.org/10.1016/j.jpainsymman.2020.03.041>
- Lee D, Choi B (2020) Policies and innovations to battle COVID-19 – a case study of South Korea. *Health Policy Technol* 9(4):587–597. <https://doi.org/10.1016/j.hlpt.2020.08.010>
- 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 Sens* 12(10):1576. <https://doi.org/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. *Environ Sci Pollut Res* 28: 6587–6599. <https://doi.org/10.1007/s11356-020-10962-2>
- McNeil LM, Kelso TS (2013) *Spatial temporal information systems : an ontological approach using STK®*. CRC Press
- Megaloiconomos PD, Thaler M, Khosravi I, Igoumenou VG, Bonanzinga T, Ostojic M, Couto AF, Diallo J (2021) Impact of the COVID-19 pandemic on orthopaedic and trauma surgery training in Europe. *Int Orthop* 44(9):1611–1619. <https://doi.org/10.1007/s00264-020-04742-3>
- Mellish TI, Luzmore NJ, Ashfaq SA (2020) 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. *J Glob Faultlines* 7(1):9–45. <https://doi.org/10.13169/jglobfaul.7.1.0009>
- Menut L, Bessagnet B, Mailler S, Pennel R, Cholokian A, Siour G (2021) Impact of lockdown measures to combat COVID-19 on air quality over Western Europe. *Sci Total Environ* 741:140426. <https://doi.org/10.1016/j.scitotenv.2020.140426>
- Moghadam SM, Shoukat A, Fitzpatrick MC, Wells CR, Sah P, Pandey A, Sachs JD, Wang Z, Meyers LA, Singer BH, Galvani AP (2020) Projecting hospital utilization during the COVID-19 outbreaks in the United States. *Proc Natl Acad Sci* 117:9122–9126. <https://doi.org/10.1073/pnas.2004064117>
- Mollalo A, Vahedi B, Rivera KM (2020) GIS-based spatial modeling of COVID-19 incidence rate in the continental United States. *Sci Total Environ*:728 <https://doi.org.libproxy.txstate.edu/10.1016/j.scitotenv>
- Nakaya T (2016) GWR4.09 UserManual, pp 2–27. <http://77pdfs.com/gwr4.09-user-manual-pdf>
- Qu, Jie-Ming, Bin Cao, and Rong-Chang Chen. 2020. COVID-19. [Electronic Resource]: The Essentials of Prevention and Treatment. Elsevier. <https://searchebcohost-com/login.aspx?direct=true&db=cat00022a&AN=txi.b5571489&site=eds-live&scope=site>.
- Ros F, Kush R, Friedman C, Zorzo EG, Corte PR, Rubin JC, Sanchez B, Stocco P, Van Houweling D (2021) Addressing the COVID-19 pandemic and future public health challenges through global collaboration and a data-driven systems approach. *Learn Health Syst* 5(1): 1–12. <https://doi.org/10.1002/lrh2.10253>
- Rosenkrantz L, Schuurman N, Bell N, Amram O (2020) The need for GIScience in mapping COVID-19. *Health Place* 67:102389.
- Sha D, Malarvizhi AS, Liu Q, Tian Y, Zhou Y, Ruan S, Dong R, Carte K, Lan H, Wang Z, Yang C (2020a) A state-level socioeconomic data collection of the United States for COVID-19 research. *Data* 5(118): 118. <https://doi.org/10.3390/data5040118>
- Sha D, Malarvizhi AS, Liu Q, Tian Y, Zhou Y, Ruan S, Dong R, Carte K, Lan H, Wang Z, Yang C (2020b) A state-level socioeconomic data collection of the United States for COVID-19 research. *Data* 5(4): 118
- Smith CD, Mennis J (2020) Incorporating geographic information science and technology in response to the COVID-19 pandemic. *Prev Chronic Dis* 17:E58. <https://doi.org/10.5888/pcd17.200246>
- Tobler, WR (1970) A computer movie simulating urban growth in the Detroit region. *Econmic Geography* 46:234–240. <https://doi.org/10.2307/143141>
- Worldometer. The United States coronavirus 2020. <https://www.worldometers.info/coronavirus/country/us/> (accessed February 28, 2021).
- Yang C, Sha D, Liu Q, Li Y, Lan H, Zhang Z, Wang Z et al (2021) Taking the pulse of COVID-19: a spatiotemporal perspective. *Int J Digital Earth* 13(10):1186–1211. <https://doi.org/10.1080/17538947.2020.1809723>
- Yuan Z, Xiao Y, Dai Z, Huang J, Zhang Z, Chen Y (2020) Modelling the effects of Wuhan’s lockdown during COVID-19, China. *Bull World Health Organ* 98(7):484–494. <https://doi.org/10.2471/BLT.20.254045>
- Zou H, Xue L (2018) A selective overview of sparse principal component analysis. *Proc IEEE, Proc IEEE* 106(8):1311–1320. <https://doi.org/10.1109/JPROC.2018.2846588>

**Publisher’s note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.