

Geospatial analysis of misinformation in COVID-19 related tweets

Amir Masoud Forati, Rina Ghose^{*}

Department of Geography, University of Wisconsin- Milwaukee, Milwaukee, WI 5321, USA

ARTICLE INFO

Keywords:

GIS
Discourse analysis
COVID-19
Coronavirus
Spatial analysis
Misinformation
Social media activity
Downplay
MGWR

ABSTRACT

COVID-19 has emerged as a global pandemic caused by its highly transmissible nature during the incubation period. In the absence of vaccination, containment is seen as the best strategy to stop virus diffusion. However, public awareness has been adversely affected by discourses in social media that have downplayed the severity of the virus and disseminated false information. This article investigates COVID-19 related Twitter activity in May and June 2020 to examine the origin and nature of misinformation and its relationship with the COVID-19 incidence rate at the state and county level. A geodatabase of all geotagged COVID-19 related tweets was compiled. Multiscale **Geographically Weighted Regression** was employed to examine the association between social media activity and the spatial variability of disease incidence. **Findings suggest that MGWR could explain 80% of the COVID-19 incidence rate variations indicating a strong spatial relationship between social media activity and spread of the Covid-19 virus.** Discourse analysis was conducted on tweets to index tweets downplaying the pandemic or disseminating misinformation. Findings indicate that sites of Twitter misinformation showed more resistance to pandemic management measures in May and June 2020 later experienced a rise in the number of cases in July.

1. Introduction

Coronavirus (COVID-19) related illness has been identified as a global pandemic by the World Health Organization (WHO). Known in the scientific communities as Severe Acute Respiratory Syndrome Coronavirus 2 or SARS-CoV-2, it was first identified in Wuhan, China in winter, 2019 (Linton et al., 2020; Xie & Chen, 2020). The virus is highly transmissible during the incubation period. Susceptibility to the virus is more common among those who have underlying conditions such as hypertension, diabetes, and heart disease (Xie & Chen, 2020). Combating the spread of virus through containment strategies was advocated by WHO and other public health agencies since early 2020s. These containment strategies (the practice of using face masks, hand-washing, social distancing, self-isolation, quarantine, enforced lockdowns) have been found to be effective in reducing transmissions, while vaccine research and deployment have been underway. As of February 23, 2021, COVID-19 has claimed 2.48 million lives globally, with USA ranking first in both the number of cases (28.2 million) and in the number of deaths (over 550,000) (Centers for Disease Control and Prevention -CDC, 2020).

Despite the high number of deaths, political and social resistance to containment measures have been profound in the USA, fueled by

misinformation spread through social media. Pollings from the early months of the pandemic have shown that many Americans are misinformed about COVID-19. Results from a poll conducted by YouGov and The Economist in early March 2020 revealed that 49% of Americans believed the Coronavirus to be manmade, 44% felt that the dangers of the virus was being exaggerated for political reasons, and 13% believed it was a hoax (Economist, 2020).

Social media has a tremendous impact on public perception and awareness of COVID-19, as 68% of American adults have reported to receiving news from social media. Further, 59% of Twitter users have found it to be an effective and reliable source for obtaining health information (Singh et al., 2020). Yet, health misinformation is a significant problem in all social media platforms, and the deluge of COVID-19 related misinformation through social media has been identified as 'infodemic' (Kulkarni, Prabhu, & Ramraj, 2020). When Anthony Fauci, Director of the National Institute of Allergy and Infectious Diseases, was asked about the impact of social media on the public's reaction to COVID-19, he responded "It [social media] has impacted it more negatively than positively. One of the problems is when disinformation gets in there, it has a way of self-propagating itself to the point where you don't know what's true and what's not true" (Gander, 2020).

This article investigates the association between social media

^{*} Corresponding author.

E-mail addresses: aforati@uwm.edu (A.M. Forati), rghose@uwm.edu (R. Ghose).

misinformation and spatial variability in the spread of the virus in USA. We focus on Twitter activity on COVID-19, as 69.3 million Americans use Twitter as their preferred social media platform. Twitter facilitates the exchange of microblogs, coined as 'tweets' - a short textual message and metadata. Geotagged tweets provide location information and are a rich source of location-based data (Sui & Goodchild, 2011). These georeferenced posts in social media (Tsou, 2015) enable us to examine the dynamic interactions between human behavior and the environment (Shaw, Tsou, & Ye, 2016). Cao et al. (2015) highlighted social media's role as a proxy to understand human behaviors and complex social dynamics in geographic spaces.

Our study examines COVID-19 related geotagged tweets during the months of May and June 2020 to investigate their relationship with the COVID-19 incidence rate at the state and county level. We use both qualitative and quantitative research methods in our study. We identify and analyze the nature of false information in Tweets through discourse analysis. We map and analyze the geographic distribution of tweets and examine the demographic characteristics of places of origin. We conduct spatiotemporal data analysis through Multiscale Geographically Weighted Regression (MGWR), to examine the association between covid-19 related Twitter activity and the spatial variability of disease incidence. Our findings indicate that such a correlation exists, necessitating the need for policy intervention.

2. Background and literature review

While COVID-19 virus had been identified in China in late 2019, its arrival in USA went undetected till January 19th, 2020 when the first COVID-19 case was confirmed in the state of Washington (Holshue et al., 2020). Thereafter, the virus spread rapidly across USA, and on March 26th, 2020, USA became the leading country in the number of cases worldwide. A year later, USA continues to lead the world in terms of both cases and deaths. Experts have identified several factors that have led to "an unnecessarily brutal pandemic, including a lack of clear messaging from the country's leadership, state and local leaders loosening restrictions too quickly, large holiday celebrations and continued resistance to wearing face masks or social distancing" (Maxouris, Yan, & Vera, 2021). A primary concern in the face of Covid-19 mitigation efforts is social unrest caused by the perceived loss of liberties arising from quarantine and shelter in place orders (Timmis & Brüssow, 2020). Relational theorists Kasapoglu and Akbal (2020) framed COVID-19 in terms of 'uncertainties' in social relations. Such uncertainties are fueled by fear caused by political, social, and economic anxieties, creating a 'moral panic' (ibid). Misinformation circulated through social media contributes significantly to such 'moral panic.' Existential fears of illness and death catastrophize thinking, further destabilizing existing social-relational structures. Therefore, it is crucial for us to examine misinformation in social media and its effects on public perception. Otherwise, misinformation will undermine global efforts to control COVID-19 virus.

The global popularity of social media usage and the ease of transmitting misinformation across the world (Ahinkorah, Ameyaw, Hagan, Seidu, & Schack, 2020) along with its potential adverse effects on public health (Lewandowsky, Ecker, & Cook, 2017) have prompted scholars to evaluate public belief in and susceptibility to COVID-19 misinformation. Bastani and Bahrami (2020) conducted a qualitative study on COVID-19 related misinformation via social media in Iran. They concluded that cultural factors, high information demand, social media prevalence, and inadequate legal supervision of online content are crucial to misinformation dissemination. Brennen, Simon, Howard, and Nielsen (2020) note that COVID-19 misinformation are presented in different forms, are gathered from various sources and make different claims; about 59% of the misinformation in their sample dataset involved various forms of reconfiguration of true information while 38% was completely fabricated.

Health misinformation is a serious problem across all social media

platforms and includes misinformation on global health crises such as Ebola outbreak in 2014, Zika outbreak in 2016 and COVID-19 (Singh et al., 2020). In their analysis of Twitter misinformation, Singh et al. (2020, 15) identified five major misinformation myths: These include "Origin of COVID-19, Vaccine Development, Flu Comparison, Heat Kills Disease, Home Remedies". Their findings show that the myth regarding the origins of the virus were highly dominant in Twitter in January and February of 2020. By the end of February, the dominant myths also included comparison of COVID-19 to flu and promotion of home remedies as cure. Other myths such as the effectiveness of heat killing COVID-19 and vaccine development theories arose over time as well. Misinformation promoting the use of Chloroquine as an effective preventive measure led to the death of an Arizona resident, who consumed a form of chloroquine used for treating aquariums (Waldrop, Alsup, & McLaughlin, 2020). In Nigeria, healthcare workers identified numerous cases of Chloroquine overdose after alleged claim from the media that it's an effective treatment of COVID-19 (Busari & Adebayo, 2020).

Conspiracy theories promoting the pandemic as a hoax or as a bio-weapon designed by sinister forces are associated with reduced containment-related behavior. By framing containment strategies as a tyrannical act of the state that violates an individual's right to freedom, conspiracy theories have championed anti containment behavior (anti-mask, anti-social distancing, anti-isolation/quarantine) as a victory of individual freedom over state's regulations, leading to faster spread of the virus (Imhoff & Lamberty, 2020). Stanley, Barr, Peters, and Seli (2020) concluded that people who are less likely to engage in effortful, deliberative, and reflective cognitive processes are more likely to believe that the pandemic was a hoax and thus less likely to engage in social-distancing and handwashing, thus accelerating the spread of the virus (Stanley et al., 2020). Uscinski et al. (2020) studied the psychological foundations of conspiracy beliefs. They noted that beliefs in conspiracies about the virus are associated with a propensity to reject information from expert authorities, consequently reducing people's willingness to comply with public health guidance. Further, racial prejudice has been associated with COVID-19 in social media discourses. Budhwani and Sun (2020) conducted research on Twitter activities to assess the prevalence and frequency of the phrases "Chinese virus" and "China virus." They found a rise in such tweets over time, suggesting that knowledge translation is likely occurring online, and COVID-19 related racist stigma is likely being perpetuated on social media.

Health misinformation can harm the public's health and is a critical threat to public health (Chou, Oh, & Klein, 2018; Seymour, Getman, Saraf, Zhang, & Kalendarian, 2015). Studies have been conducted to investigate the association between COVID-19 misinformation and public health guidance compliance. Bertin, Nera, and Delouvée (2020) conducted two cross-sectional studies exploring the relationship between COVID-19 misinformation and attitudes towards vaccines as well as support towards controversial medical treatment such as chloroquine. They noted that COVID-19 conspiracy beliefs, as well as a conspiracy mentality, are associated with a distrust of vaccines and negatively related to participants' intentions to be vaccinated against COVID-19 in the future. Roozenbeek et al. (2020) investigated how susceptibility to misinformation about Covid-19 affects key self-reported health behaviors. They demonstrate a clear association between susceptibility to misinformation and vaccine hesitancy and, therefore, a reduced likelihood of complying with public health guidance, causing a rapid spread of the virus (Krause, Freiling, Beets, & Brossard, 2020). In South Korea, the results of a cross-sectional online survey shows that higher exposure to COVID-19 misinformation led to a decline in individual preventive behaviors; the study highlights the potential of misinformation to undermine global efforts in COVID-19 disease control and faster and deeper disease spread (Lee et al., 2020). Delays, denials, and misinformation about COVID-19 have exacerbated the virus spread and slowed pandemic response (Abutaleb, Dawsey, Nakashima, & Miller, 2020).

Many factors affect the framing and spread of health misinformation. Törnberg (2018) finds that the spread of misinformation varies by

subject and that some demographic groups and cultures may be more vulnerable to misinformation than others. Misinformation is particularly harmful to marginalized/vulnerable people, who live in more precarious circumstances. Individuals can also be subjected to various forms of disinformation depending on their information-seeking patterns (Golebiewski & Boyd, 2018; Marwick & Lewis, 2017; Nsoesie, Oladeji, & Sengeh, 2020). This is particularly true for the novel COVID-19, of which little was known.

Conspiracy theories, pessimistic attitudes toward science, and political philosophy are all linked to individuals' interrelated dispositions. Guess, Nagler, and Tucker (2019) note that people who are older and more politically conservative tend to post more political misinformation online. Individuals that gravitate toward conspiracy theories on social media are often more likely to engage with disinformation arguments (likes, comment, share) than those who gravitate toward scientific narratives (Bessi et al., 2015). To formulate a socio-cognitive profile of individuals likely to spread misinformation online, Lobato, Powell, Padilla, and Holbrook (2020) conducted an exploratory survey to investigate individuals' willingness to distribute COVID-19 misinformation over social media-the study examined their proclivity towards conspiracy theories, their views on science, and their political philosophy. Their results suggest that political belief, particularly social dominance orientation, has an impact on tendencies to share different kinds of misinformation and conspiracy theories.

The ability to access health records on the internet has changed the way people learn about their illnesses. While the internet has a wealth of useful information, many health Web pages contain inaccurate or misleading information. Misinformation about health can have a wide range of consequences in people's lives. Misinformation may influence people's perceptions of a disease's impact, effective preventive behaviors, and even memories of their own past experiences (Greenspan & Loftus, 2021). Offering digital skills and information literacy education to digitally disadvantaged groups is of paramount importance, as topic involvement, self-efficacy, digital media use, age, and education have an impact on misinformation vulnerability (Chesser, Burke, Reyes, & Rohrberg, 2016; Jaeger, Bertot, Thompson, Katz, & DeCoster, 2012; Seo, Erba, Altschwager, & Geana, 2019). Benotsch, Kalichman, and Weinhardt (2004) surveyed 324 adults with HIV to examine their usage of Internet for obtaining health information. Their results show that vulnerability to misinformation is directly related to lower income and lower educational attainment, poorer reading comprehension, lower literacy level, and irrational health beliefs. Seo, Blomberg, Altschwager, and Vu (2020) examined how low-income African American older adults assess the credibility of online information. They note that education and topic involvement are statistically significant factors associated with assessments of message content and source credibility and has an impact on misinformation vulnerability.

Our research focuses on the impact of misinformation spread through Twitter, as "59% of Twitter users have reported it as good or extremely good in sharing preventive health information" (Singh et al., 2020, p. 2). While examining the spread of true and false news stories through Twitter, Vosoughi, Roy, and Aral (2018) found that falsehood diffuses significantly farther, faster, deeper, and more broadly than the truth. The novelty element of false news generates greater public interest over truthful news accounts; consequently, false news is rapidly spread as people are more likely to share novel information (Vosoughi et al., 2018). Yang, Torres-Lugo, and Menczer (2020) studied the extent of links to low credibility information on Twitter during the pandemic and contend that the combined volume of tweets linking to low-credibility information is comparable to the volume of New York Times articles and CDC links, raising concern about the volume and extent of low credibility information related to COVID-19 on Twitter. Brennen et al. (2020) note that prominent public figures play a significant role in spreading misinformation about COVID-19 and that there are significant motivations behind such activities. It is imperative that trusted fact-checkers, social media activists, and media organizations

continue to hold prominent figures to account for claims they make on social media. Pennycook, McPhetres, Zhang, Lu, and Rand (2020) suggest that social bots are more likely to be involved in posting and amplifying low-credibility information and called for future research to investigate the impacts of misinformation spread by social bots upon COVID-19 incidence rates at the state and county level. Kouzy et al. (2020) examined a sample of 673 tweets to understand the impacts of misinformation on public health. Their findings indicate that medical misinformation about the COVID-19 epidemic is being propagated at an alarming rate on social media. Singh et al. (2020) note that a strong spatiotemporal relationship exists between Twitter information flow and new cases of COVID-19. They note that Twitter conversations around COVID-19 myths led to an increase in COVID-19 cases by 2–3 days. All in all, COVID-19 misinformation (i.e., misleading news content) and rumors are propagating at an alarming pace, masking other credible healthy behaviors like hand washing, social distancing, and promoting incorrect practices that will potentially increase the virus spread. Such misinformation is drowning official public health advice on COVID-19, making it extremely problematic for healthcare professionals' voices to be heard, the implications of which may be enormous as the virus spreads faster and deeper (Oxford Analytica, 2020).

3. Study site, data and methodology

Informed by the literature on social media misinformation and its effects upon the spread of COVID-19, this paper examines the relationship between Twitter misinformation and COVID-19 incidence rate in USA. This study examines Twitter misinformation in the months of May and June 2020. We studied 84,864 geotagged tweets on COVID-19 to identify those that originated in USA. In total, we found 37,587 COVID-19 related tweets originating from USA. We used the Twitter Streaming API for data gathering, as it enables the collection of data using tweets' texts, hashtags, user information, and location. To gather tweets based on selected hashtags related to COVID-19, we used the Tweepy API (Roesslein, 2009) to collect all geotagged tweets written in the English language from May 1st to June 30th, 2020. We filtered the data by using trending hashtags such as "corona", "coronavirus", "COVID", "pandemic", "lockdown", "quarantine", "hand sanitizer", "ppe", "n95", "sarscov2", "nCov", "COVID-19", "ncov2019", "2019ncov", "flatteningthecurve", "social distancing", "workfromhome" and the respective hashtag of all these keywords using filter: language "English".

To understand the types and nature of misinformation, we used qualitative research methods to extract and analyze tweets. Qualitative methods provide an opportunity to conduct textual analysis or discourse analysis of individual tweets, which helps to understand the practices of a specific user group. We used discourse analysis to examine 37,587 geotagged tweets to identify tweets that (i) downplayed COVID-19, (ii) showed resistance toward safety measures, (iii) disseminated COVID-19 conspiracy theories (iv) propagated misinformation. To examine the demographic characteristics of tweeters, we obtained demographic data from the U.S. Census Bureau. We collected COVID-19 data at the county level from USAFacts (usafacts.org). Next, we computed and joined case numbers to the administrative boundary shapefile of counties obtained from the TIGER/Line database (www.census.gov), using ArcGIS Desktop 10.7.

Spatial statistics and visualization helped us to statistically investigate the geographic relationship/correlation between several explanatory variables and disease outbreak (Watkins et al., 2007; Wang, Minnis, Belant, & Wax, 2010; Dom, Ahmad, Ishak, & Ismail, 2013). Therefore, we conducted multiscale spatial modeling to (i) measure the spatial scale at which Twitter activity operates based on the covid-19 pandemic incidence rate, (ii) examine the ways that individuals react to and shape the built environment, (iii) capture the spatiotemporal patterns of tweets, (iv) examine its spatial heterogeneity. We conducted partial distance correlation to examine the correlation between social media activity and the COVID-19 incidence rate. Distance correlation is a

measure of **dependence between random vectors**. The population distance correlation coefficient is zero if and only if the random vectors are independent. Therefore, distance correlation measures both **linear and nonlinear correlation between two random variables**. This contrasts with Pearson's correlation, which can only detect the linear association between two random variables (Székely, Rizzo, & Bakirov, 2007). Partial distance correlation measures the correlation between two random variables, while the effect of a set of controlling random variables is eliminated (Székely & Rizzo, 2014). We calculated partial distance correlation coefficient between two variables: the number of tweets and the number of confirmed cases, while the effect of a controlling random variable population is removed to investigate both linear and nonlinear correlation between these variables.

We used Poisson Multiscale Geographically Weighted Regression (MGWR) as our spatial modeling tool, as it can explain COVID-19 variation at multiple scales (Mollalo, Vahedi, & Rivera, 2020). Brunson, Fotheringham, and Charlton (1996) introduced Geographically Weighted Regression (GWR) as an extension of global traditional regression models to allow for spatially varying variables so that variables can be derived for each location separately. GWR is a method that is recently employed to understand how spatial processes associated with crises and crises themselves vary across space (Chun, Chi, & Hwang, 2017; Purwaningsih, Prajaningrum, & Anugrahwati, 2018; Rifat & Liu, 2020). GWR adjusts for nonstationarity in relationships by the use of a data-borrowing procedure in order to perform a series of local regressions for each area, which enable us to estimate model's parameters at any given locations in a study area, in contrast to a traditional "global" ordinary least squares (OLS) regression model that estimates a single set of parameters that are presumed to be constant in the study area. Comparison of local parameter estimates across space is beneficial because it helps to examine any variability in social media activity across geographic space, which is overlooked in a global model. Therefore, GWR offers a mechanism not only to explore whether a model is an appropriate representation, but also to define the variables that lead to such a representation for specific locations (Fotheringham, Yang, & Kang, 2017; Oshan, Li, Kang, Wolf, & Fotheringham, 2019). GWR assumes that each determinant operates at the same spatial scale or same kernel bandwidth. Nevertheless, it is much more likely that the complex issue of social media activity associated with the Covid-19 pandemic may vary at different scales or unique kernel bandwidths (Oshan et al., 2019). When it is presumed that the same spatial scale applies to all relationships, it is likely that the true trends across space are distorted since the model is misspecified. Consequently, when modeling complex spatial processes, it is important to use a multiscale approach. In summary, MGWR is an extension of GWR that allows us to examine relationships at varying spatial scales, by deriving an optimal bandwidth vector in which each element indicates the spatial scale at which a particular process takes place as opposed to a single, constant bandwidth for the entire study area.

MGWR can be calculated in the following manner (Fotheringham et al., 2017):

$$y_i = \sum_{j=0}^m \beta_{bwj} X_{ij} + \varepsilon_i, \quad i = 1, 2, \dots, n$$

Where at county i , y_i is the number of COVID-19 cases, X_{ij} is the vector of explanatory variables (number of tweets and population), β_{bwj} is the bandwidth used for calibration of the j th relationship, and ε_i is a random error term (Fotheringham et al., 2017).

We use Poisson MGWR to investigate the relationship between the potential explanatory variables (counties' number of tweets per 1000 residents and digital divide scores) and the dependent variable (COVID-19 incidence rate per county). An (adaptive) bi-square kernel, which removes the effect of observations outside the neighborhood specified with the bandwidth and (minimize) correct Akaike Information Criterion (AICc), was used to select optimal bandwidth (Oshan et al., 2019).

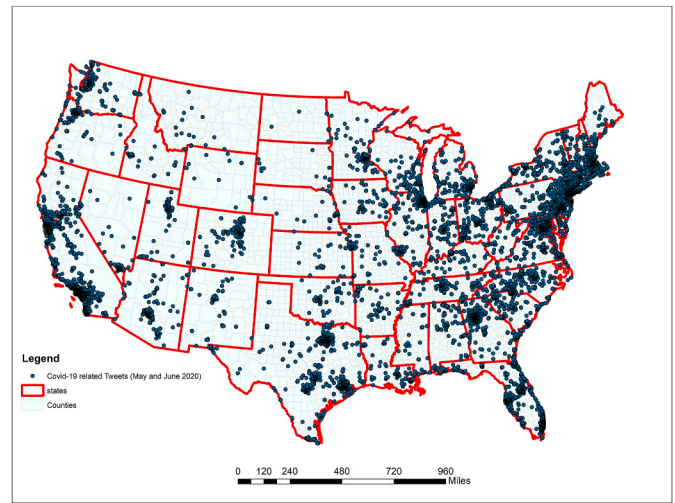


Fig. 1. Distribution of Covid-19 related tweets in May and June 2020, throughout conterminous United States.

Table 1

Top ten counties with the highest number of COVID-19 related geotagged tweets in May and June.

County	State	No. of geotagged tweets in May and June	No. of confirmed cases as of July 1st	No. of confirmed deaths as of July 1st
Los Angeles	California	3290	105507	3402
New York	New York	3075	28518	3088
Hudson	New Jersey	1541	18842	1457
Fulton	Georgia	724	7444	314
Queens	New York	719	65455	7059
Cook	Illinois	664	90911	4581
Harris	Texas	658	31422	378
Travis	Texas	623	9527	124
Kings	New York	615	59507	7104
Riverside	California	590	18041	463

We used MGWR V 2.2.1 as our modeling tool (<https://sgsup.asu.edu/sparc/multiscale-gwr>).

4. Results and discussion

4.1. Examining the distribution and content of COVID-19 tweets

As the first step in our analysis, we examined the geographic distribution of the extracted 37,587 COVID-19 related tweets. The following figure (Fig. 1) shows the distribution of geotagged tweets throughout the conterminous USA.

The following table (Table 1) shows the top ten counties in terms of the number of COVID-19 related geotagged tweets in May and June, 2020.

To understand the nature of false information and their geographic distribution, we examined the content of each tweet and assigned it an index. Out of 37,587 tweets, 821 tweets contained sentiments that we

Table 2

Most popular hashtags among tweets sharing misinformation.

Pandemic	CovidPropaganda	NoMask
EndTheShutdown	DefundTheFDA	MaskFree
BiologicalWarfare	HealthFreedom	Scamdemic
DontStayAtHome	GovernmentControl	CoronaBS

identified as negative – tweets that (i) downplayed the severity of COVID-19, (ii) propagated conspiracy theories, (iii) disseminated false news/facts about COVID-19. We also examined hashtags used in these tweets. Besides the #covidiot, which has been used by both indexed and unindexed tweets, the most popular hashtags are used by indexed tweets against containment measures, downplaying the virus, or disseminate misinformation is presented in the following table (Table 2).

Below we provide some examples of tweets using these hashtags:

"Welp, add another casualty to the Plandemic, sweettomatoes / soup plantation is officially done for. They are closing all locations permanently. They are closing all 97 locations permanently in wake of the Corona"

"REVOLUTION -Careful- this dangerous - true and proven research will GET you banned and censored from every social media platform: After studying global data from the novel Coronavirus (COVID-19) pandemic, researchers have discovered a strong correlation between severe vitamin D deficiency and mortality rates. #healthmediastar #immunity #healthmedia #healthygut #healthygirl #healthyvegan #veganlifestyle #veganinspiration #naturopath #naturopathicrevolution #vitamnind #lockdownlife #lockdown2020 #quantum"

"Lies they tell??? #lies #covid19 #governmentcontrol #riotworldwide #raisethefire #conjuredhoodoo #nlhealngtempal".

We found that tweets containing such hashtags primarily originated from Orlando, FL, Dallas, TX, Palm Beach, FL, Houston, TX, Los Angeles, CA, and Watchung, NJ. The cities with the highest use of these hashtags in May and June are located in Texas, Florida, and California, which experienced massive surges in COVID-19 cases in July and are emerging as the epicenters of pandemic after reporting record numbers of new confirmed cases for weeks in a row (Hawkins, Kornfield, Copeland, & Sonmez, 2020).

We next aimed to categorize tweets containing false information based on content analysis. Based on our content analysis, we categorized all indexed tweets into two groups: those that provide misinformation and those that downplay the severity of the virus. The first category includes 45.33% of tweets which propagated false information about the effectiveness of containment measures. Further, these tweets claimed that religious faith is sufficient protection as God will protect His devout followers from the virus. Alcohol consumption was also promoted as an effective preventive measure from COVID-19 prevention. Tweets also championed various conspiracy theories. One theme framed the COVID-19 pandemic as a hoax created by government to control people and limit their freedom. Anti-state and anti-regulatory in nature, these tweets actively promote violation of containment strategies as an act of individual liberty. Another theme framed COVID-19 as a bio-engineered virus or biological weapon created by sinister political powers. Examples below illustrate the nature of these tweets:

"This came from the CDC website. The 65k deaths claimed to b from Coronavirus is wrong. The actual death number is 11k. So the lockdown everyone is suffering from is based on lies"

"More tyranny from Dr Fauci. #Covid #EndTheShutdown Where are the lawyers when you need them?"

"Not letting the Covid Communists stop me from seeing my friends and having a good time. Don't be afraid to have a good time.) @ Los Angeles, California"

The second category includes 55.64% of tweets that downplayed the severity of Coronavirus, the process of transmission and lifetime of the virus. Examples are provided below:

"I come back the covid didn't kill me ... à West Hollywood, California"

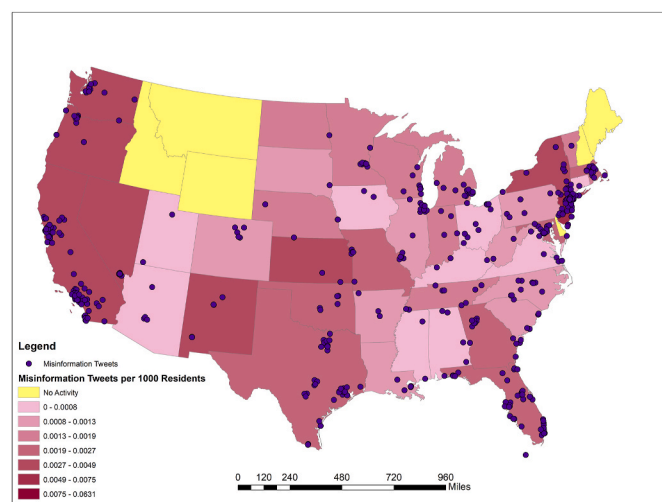


Fig. 2. Distribution of Misinformation Tweets in May and June 2020, throughout conterminous United States.

Table 3

Top 10 states leading in misinformation tweets.

Rank	State	Flagged tweets per 1000 residents
1	District of Columbia	0.0632
2	New Jersey	0.0075
3	Kansas	0.0049
4	New York	0.0047
5	California	0.0045
6	Oregon	0.0044
7	Washington	0.0038
8	Hawaii	0.0036
9	New Mexico	0.0034
10	Nevada	0.0033

"I had COVID without symptoms. I have IgG antibodies, I have no virus. Immunoglobulin G provided immunity that can help others. #sars #coronavirus #covidart #covidfitness #recovered #sarscov2 @ Williamsburg, Brooklyn"

This category of tweets promoted anti containment behavior by minimizing the severity of COVID-19. Consequently, both categories of tweets influenced individuals to reject containment measures, leading to further transmission and infections.

Next, we computed the numbers of misinformation tweets per county and state and mapped the results. Fig. 2 shows the distribution of misinformation tweets throughout the conterminous USA.

Table 3 shows the list of states with the highest number of indexed tweets per 1000 persons.

Notably, five of these top 10 states: Kansas, Oregon, California, Washington, Hawaii, and Nevada, have experienced a massive surge in their known COVID-19 cases as of July 1st (Hawkins et al., 2020), right after the two month case study period of this study on the extent of misinformation on Covid-19 in Twitter. Our Twitter discourse analysis findings indicate that resistance to containment measures and lack of public awareness bear a significant impact on the patterns of illness and death during the pandemic. The relatively strong association between the number of indexed tweets and the number of cases per capita supports the notion that the sites of misinformation are now experiencing a higher surge of COVID 19 cases. Dowd et al. (2020) emphasized the importance of considering population dynamics and demographic data to mitigate the approaches to combat the pandemic. A strong relationship between population, social media activity, and COVID-19 incidence rates, as our results suggest, highlights the importance of social media monitoring during the COVID-19 pandemic.

Table 4

Top 10 Counties with the highest number of tweets sharing misinformation per capita.

Rank	County	State	Population per SQMI	White (%)	Median age	Obtained Bachelor's degree or higher	Below poverty level (%)	Unemployment rate	Median income (dollars)
1	Norton	Kansas	6.50	92.97%	43.30	11.04%	12.70	1.20	49891
2	Montgomery	Kansas	53.70	84.45%	39.90	11.57%	18.30	5.50	45173
3	Grant	New Mexico	7.6	83.20%	45.8	28.81%	21.8	7.6	56094
4	Box Butte	Nebraska	10.60	88.75%	41.30	11.92%	11.90	5.20	56412
5	Poquoson	Virginia	788.90	94.09%	43.30	29.37%	4.50	5.30	96831
6	Martin	Kentucky	55.80	92.28%	37.20	5.54%	26.30	13.70	35125
7	Hudson	New Jersey	13808.8	52.70%	34.3	42.27%	16.3	6.1	97596
8	Okmulgee	Oklahoma	56.80	66.14%	38.80	9.13%	20.00	9.50	42175
9	Bremer	Iowa	56.90	94.41%	39.30	19.81%	8.20	3.20	68023
10	Tillamook	Oregon	22.9	91.65%	47.4	21.81%	15	4.8	63543

Next, we examined the data at the county level to identify the top ten counties with the highest number of indexed tweets per capita. These are Norton (Kansas), Montgomery (Kansas), Grant (New Mexico), Box Butte (Nebraska), Poquoson (Virginia), Martin (Kentucky), Hudson (New Jersey), Okmulgee (Oklahoma), Bremer (Iowa), and Tillamook (Oregon). The number of COVID-19 cases from May and June to early July has almost doubled (1.88 times) in these counties as well.

To examine the demographic context behind such tweets, we used census data in conjunction with our geolocated Twitter data set. Table 4 summarizes the key demographic characteristics of the residents in these counties (www.data.census.gov).

The demographic data associated with the geotagged tweets indicate the following characteristics: The Twitter users are predominantly White (84% on average) and relatively middle-aged (average median age is 41 years old), middle-income (average median income is 61,086), and just about 19% of the population 25 years old and over have obtained a Bachelor's degree or higher. Most of these counties are categorized as rural with extremely low population density (Office of Rural Health Policy, 2015). Voting patterns from 2016 indicate a conservative mindset among the Twitter users, as, on average, 62.13% voted for the GOP party (Presidential Election Results: Donald J. Trump Wins, 2016). Lower educational attainment may be a contributing factor in the spread of misinformation. A lack of understanding of epidemiology and scientific medical research can lead to a rejection of scientific knowledge. Low population density can lead to feelings of isolation, which can be mitigated through social media participation. The desire to gain followings by spreading novel and sensationalistic information is also likely a contributing factor behind such Twitter activities. This is particularly damaging as many rural counties either have no intensive care units in their hospitals or no hospitals at all (Ajilore, 2020). Finally, partisan politics may have shaped the attitude of Twitter users. The relationship between Twitter activities that provide misinformation/downplay the significance of the virus has, in turn, affected the spread of COVID-19 (Singh et al., 2020). Therefore, it is imperative to raise public awareness of scientific knowledge and block fake remedies, myths, and false news about COVID-19. The impact of the digital divide on the lack of public awareness must also be considered. During a pandemic, health officials rely on the Internet and social media sites (and other digital platforms) to communicate vital information to the public. However, the effectiveness of these digital channels depends on whether individuals have access to it. Thus, concerns have been raised about the digital divide, information quality, and biases (Goodchild & Li, 2012; Oh, Kwon, & Rao, 2010), as well as source credibility (Ostermann & Spinsanti, 2011). Recent literature suggests that demographic groups (i.e., low income, low education, and elderly populations) may lack the resources, skills, and motivations to access social media, and therefore, they may be less likely to post relevant information through social media (Xiao & Huang 2015). Sui, Goodchild, and Elwood (2013) report that two-thirds of humanity does not have access to the rapidly expanding digital world. At least 10% of the US population does not use the internet (Anderson, Perrin, Jiang, & Kumar, 2019). Therefore, we

must recognize that user-generated data will provide only selective representations of any issue and that there will always be people and communities that are missing from the map (Zook, Graham, Shelton, & Gorman, 2010; Burns, 2015; Meier, 2012; Ziemke, 2012). Thus, this study's result should be seen not as reflections of on-the-ground conditions but instead as a representational negotiation rooted in spatial inequalities. The impacts of the digital divide upon social media usage are beyond the scope of this paper but should be considered in future research.

4.2. Examining correlations between twitter activity and COVID-19 incidence rate

To determine correlation between twitter activity and COVID-19 incidence rate, we undertook the following steps. First, we calculated the partial distance correlation between COVID-19 Twitter activity and the number of confirmed cases in May and June 2020, while controlling the effect of the population as a third variable. Partial distance correlation measures association between two random variables with respect to a third random variable, analogous to, but more general than (linear) partial correlation (Székely & Rizzo, 2016). In the next table (Table 2), estimated distance correlations are presented. As expected, there is a relatively weak positive correlation (0.27) between the number of tweets per county and the number of cases, indicating that the higher the number of confirmed COVID-19 cases in a county, the more geotagged COVID-19 tweets originate in that county. Therefore, people who are worse affected by the pandemic are relatively more active on Twitter.

Next, we examined the influence of digital divide. The Digital Divide Index (DDI) was proposed by Gallardo (2017) to quantify internet physical access/adoption and socio-economic characteristics that affect user motivation, skill, and usage. The DDI ranges from 0 to 100 in value, where 100 indicates the highest level of digital divide. It consists of two scores, which both range from 0 to 100: the infrastructure/adoption score (INFA) and the socio-economic score (S.E.). Five variables related to broadband internet infrastructure and adoption are grouped together in the INFA score: (1) percentage population without access to fixed broadband; (2) percentage of homes without a computing device; (3) percentage of homes with no internet access; (4) median maximum advertised internet download speeds; and (5) median maximum advertised internet upload speeds. The S.E. score combines four factors that are considered to influence technology adoption: (1) the percentage of the population aged 65 and older; (2) the percentage of the population aged 25 and up with less than a high school diploma; (3) individual poverty rate; and (4) the percentage of the population with a disability. To determine the overall DDI score, these two scores are combined (Gallardo, 2017). Using Gallardo's proposed formula, we calculated the INFA and SE scores for all census tracts in United states to be included in the modeling process as control variables.

In order to investigate the calculated correlation at different scales, we employed MGWR to examine the relationship between the explanatory variable (normalized total number of tweets per county) and the

Table 5

Partial distance correlation coefficients.

Partial distance correlation between the number of tweets and the number of confirmed cases, controlling the effect of population	0.267322
distance correlation between the number of tweets and population	0.8156
distance correlation between the number of tweets and the number of confirmed cases	0.7378
distance correlation between the number of confirmed cases and population	0.7878

Table 6

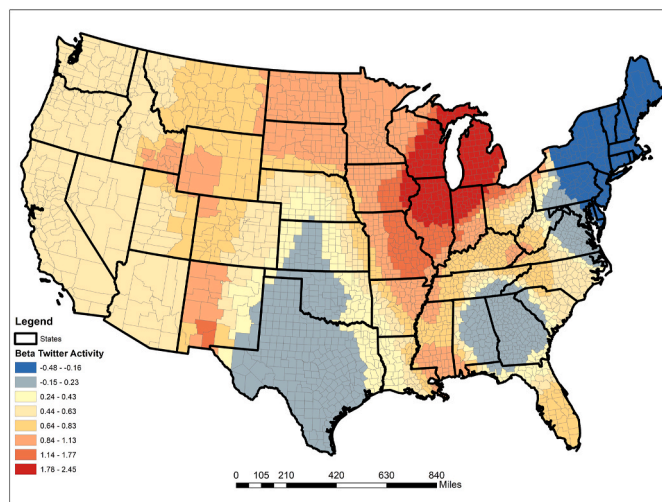
Global regression results.

Variable	Estimate	SE	p-value
Intercept	0.000	0.014	1.000
Tweet rate	0.350	0.018	0.000
SE Score	-0.183	0.023	0.000
INFA Score	-0.043	0.019	0.000

Table 7

MGWR results.

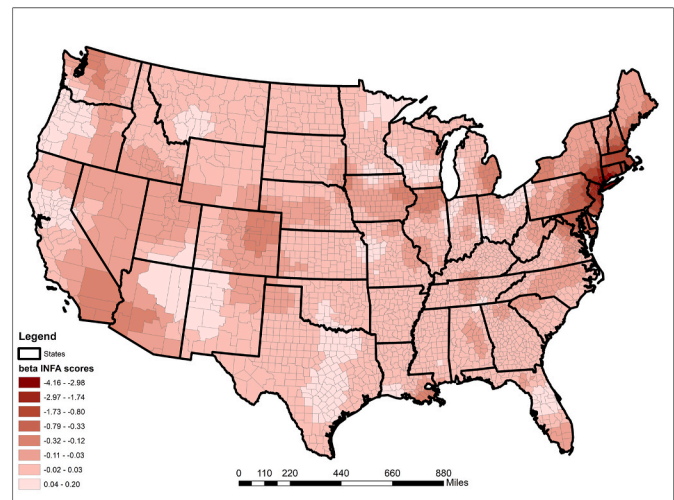
Variable	Bandwidth	Effective # Params	Critical t-value
Intercept	842	209.837	3.679
Tweet rate	242	18.146	2.996
SE Score	2171	7.667	2.007
INFA Score	44	194.236	3.659

**Fig. 3.** Composite map of MGWR parameter estimate surfaces for Tweet rate.

dependent variables (normalized COVID-19 number of confirmed cases per county, INFA score per census tracts, and SE score per census tracts).

Results from the global regression model for 3142 observations (counties) with an R-Squared of 0.424 are summarized and presented in Table 5, to provide context for the MGWR results.

The global model produces a relatively Moderate R^2 , indicating about 42% of the variation across covid-19 incidence rate can be accounted for by social media activity in this study. Based on a 95% confidence level standard t-value threshold of 1.96, the tweet rate, INFA score and SE score are statistically significant. The global regression model results assume that the relationships are constant across the study area. In order to relax this assumption, deter unexplainable high level of spatial heterogeneity, local multicollinearity, and concavity in the local subsets of the data (Oshan & Fotheringham, 2018; Oshan, Smith, & Fotheringham, 2020), and allow the processes to vary at different scales, it is necessary to employ MGWR. Calibrating an MGWR model produces

**Fig. 4.** Composite map of MGWR parameter estimate surfaces for INFA score.

a vector of optimal bandwidths that describe the spatial scale at which each process in the model varies (Oshan et al., 2020). MGWR was applied to the same set of the explanatory variable used in the global model, The R^2 increased to 0.795 in the MGWR model from 0.424 in the global model, and the AIC decreased to 4430.295 in the MGWR model from 7190.708 in the global model. MGWR model obtained a high R^2 (0.795), indicating that Twitter activity could explain about 80% of the total variations of COVID-19 incidence rates. In Table 6, the bandwidth related to the explanatory variable tweet rate is listed (theoretically, the global model assumes bandwidth of infinity). Association between covid-19 incidence rate and social media activity seems to occur at a regional scale, with bandwidth implying several hundred nearest neighbors (Table 7).

As expected, the Tweet rate has a positive non-zero parameter estimate and displays regional spatial variation (Fig. 3). The tweet rate surface highlights significant role of twitter activity in explaining covid-19 incidence rate, it is clustered all over the country except for New England Texas, and Georgia, and the association is the strongest in midwestern and central states; the characterization of this cluster requires further investigation, but it is in agreement with the MGWR local R-squared. Spatial heterogeneity in the parameter estimate identifies hot spots of high and low Covid-19 incidence rates after controlling for the Twitter activity. These spatial patterns may include both the effect of geography and the effect of geographic patterning associated with potential omitted variables, as Coffee, Lockwood, Rossini, Niyonsenga, and McGreal (2020) highlight the role of location and spatial context as an essential component in shaping human behavior and socioeconomic status.

As shown in Table 6, the socioeconomic score in MGWR is statistically non-zero and occurs at a regional scale with bandwidths implying nearest states; the infrastructure/adoption score in MGWR is statistically non-zero and occurs at a local scale with bandwidths implying nearest neighbors, suggesting the impacts of rural/urban divide in high-speed residential Internet access (Fig. 4). As the residential Internet access moves toward high-speed connections, a rural-urban divide has been created in the U.S. regarding high-speed access. Rural-urban differences in people, place, and infrastructure are all possible causes of this high-speed digital divide. This highspeed digital divide is shaped by uneven geographic supply of internet infrastructure, demographic characteristics of users (age, income, educational attainment) and in network externalities (Greenstein, 2020; Whitacre & Mills, 2007). High speed internet infrastructure tends not to be available in low-density regions, while some areas lack any internet infrastructure (Forman, Goldfarb, & Greenstein, 2018). Existence of significant intercept with regional variation is an indication of local scale determinant alongside social

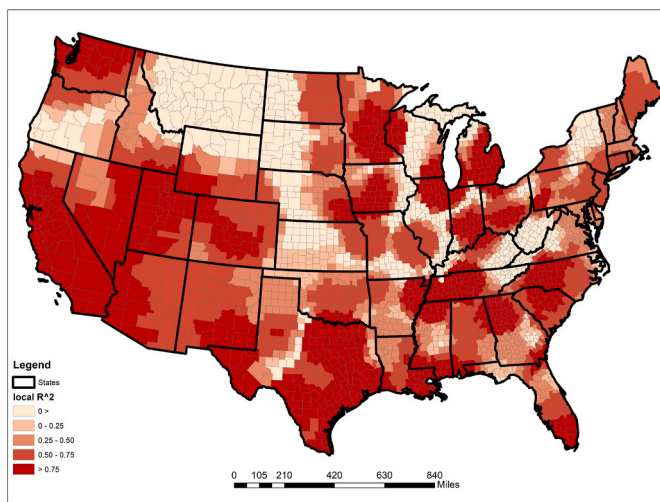


Fig. 5. Mgwr R^2 .

media activity to explain Covid-19 incidence rate. Disadvantaged groups are less likely to post crisis relevant information through social media (Xiao & Huang 2015); additionally, language barrier may affect marginalized communities from using information (Crawford & Finn, 2015); different local and state governments policies might have led to regional variation of intercept. Thus, consideration of these determinants and follow-up investigations at a local scale is necessary for future multiscale analyses.

Fig. 5 illustrates the spatial distributions of local R^2 values in the MGWR model. Several counties in Florida, California, the Tristate area, Texas, Michigan, Minnesota, Nevada, Washington, Carolinas, Utah and Louisiana show very high local R^2 , indicating a strong performance of the model in those counties. In contrast, the local R^2 values were low in most of the counties in Montana, North Dakota, South Dakota, part of Wyoming, Kansas, and West Virginia, indicating a poor performance of the model across these states. R^2 values distribution clearly shows that model works with high degree of accuracy in urban centers, coastal areas and cities while social media activity might not be a good explanatory variable to model covid-19 incidence rate in rural areas.

These findings provide explanations for the high variability of the disease incidence in the continental United States. MGWR parameter surfaces indicate that there is a strong correlation between COVID-19 cases and Twitter conversations. Our findings are thus in line with that of Singh et al. (2020). Continued monitoring of social media activity can assist in understanding the dynamics of disease spread. However, the predictive ability of the model is somewhat limited by the granularity of the data. The finest spatial granularity at which nationwide COVID-19 data is provided is at the county-level. Therefore, it is difficult to make inferences at the sub-county and individual levels accurately.

5. Conclusion

Social media conversations have a significant impact on public perception of COVID-19. Spatial modeling through MGWR explains 80% of the variability in the number of known COVID-19 cases in each county (based on Twitter activity and population), suggesting a strong association between these variables. Therefore, in the absence of other reliable indicators, analysis of Twitter conversations can help predict the spread and outbreak of COVID-19. Our research findings show how much, where, and how people are communicating about the COVID-19 pandemic. As the pandemic continues to impact people on a personal level, COVID-19 related tweets will continue to increase since people tend to care about the news that affects them personally (Singh et al., 2020). Likewise, falsehood diffuses significantly farther, faster, deeper, and more broadly than the truth (Vosoughi et al., 2018), and it is

imperative that we examine misinformation among those who are most impacted. Findings from our discourse analysis suggest that the six states of Kansas, Oregon, California, Washington, Hawaii, and Nevada, where people downplayed COVID-19 and propagated misinformation in May and June the most, have experienced a massive surge in their known COVID-19 cases as of early July. Our findings highlight the substantial impact of misinformation and resistance to containment measures on the patterns of illness and death during the pandemic. However, our study has several limitations. First, the decision to examine only geo-tagged tweets meant that we examined only a subsection of all tweets. Second, our findings are constrained by data granularity issues. We aim to address these limitations in future research. Meanwhile, social media narratives containing misinformation and misinformation on disease spread have a significant impact on public awareness. A sustained and coordinated effort by fact-checkers, social media platforms, independent media, news agencies, and public authorities is needed to control the spread of misinformation about COVID-19 so that scientific information is effectively communicated, and public awareness of the pandemic is raised.

Acknowledgment

We would like to thank anonymous reviewers for taking the time and effort to review the manuscript. Amir Masoud Forati would like to thank Rachel Hansen, Kimberly Singer, Shayan Haseli, Hamed Saadabadi, and Mousa Javidialsaadi for their contribution to discourse analysis and their help. This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

References

- Abutaleb, Y., Dawsey, J., Nakashima, E., & Miller, G. (2020). *The U.S. Was Beset by Denial and Dysfunction as the Coronavirus Raged*. The Washington Post. April 4. Retrieved from <https://www.washingtonpost.com/national-security/2020/04/04/coronavirus-government-dysfunction/?arc404=true>.
- Ahinkorah, B. O., Ameyaw, E. K., Hagan, J. E., Jr., Seidu, A. A., & Schack, T. (2020). Rising above misinformation or fake news in Africa: Another strategy to control COVID-19 spread. *Frontiers in Communication*, 5, 45.
- Ajilore, O. (2020, July 17th). Rural America is starting to feel the impact of the coronavirus. Retrieved from center for American progress <https://www.americanprogress.org/issues/economy/reports/2020/04/28/484016/rural-america-starting-feel-impact-coronavirus/>.
- Anderson, M., Perrin, A., Jiang, J., & Kumar, M. (2019). "10% of Americans don't use the internet. Who are they?" pew research center. April 22 <https://www.pewresearch.org/fact-tank/2019/04/22/some-americans-dont-use-the-internet-who-are-they/>.
- Bastani, P., & Bahrami, M. A. (2020). COVID-19 related misinformation on social media: A qualitative study from Iran. *Journal of Medical Internet Research*. <https://doi.org/10.2196/18932>
- Benotch, E. G., Kalichman, S., & Weinhardt, L. S. (2004). HIV-AIDS patients' evaluation of health information on the internet: The digital divide and vulnerability to fraudulent claims. *Journal of Consulting and Clinical Psychology*, 72(6), 1004.
- Bertin, P., Nera, K., & Delouvé, S. (2020). Conspiracy beliefs, chloroquine, and the rejection of vaccination: A conceptual replication-extension in the COVID-19 pandemic context.
- Bessi, A., Coletto, M., Davidescu, G. A., Scala, A., Caldarelli, G., & Quattrociocchi, W. (2015). Science vs conspiracy: Collective narratives in the age of misinformation. *PLoS One*, 10(2), Article e0118093.
- Brennen, J. S., Simon, F. M., Howard, P. N., & Nielsen, R. K. (2020). *Types, sources, and claims of Covid-19 misinformation*. Reuters Institute.
- Brunsdon, C., Fotheringham, A. S., & Charlton, M. E. (1996). Geographically weighted regression: A method for exploring spatial nonstationarity. *Geographical Analysis*, 28(4), 281–298. <https://doi.org/10.1111/j.1538-4632.1996.tb00936.x>
- Budhwani, H., & Sun, R. (2020). Creating COVID-19 stigma by referencing the novel coronavirus as the "Chinese virus" on twitter: Quantitative analysis of social media data. *Journal of Medical Internet Research*, 22(5), Article e19301. <https://doi.org/10.2196/19301>
- Burns, R. (2015). Rethinking big data in digital humanitarianism: Practices, epistemologies, and social relations. *Geojournal*, 80(4), 477–490. <https://doi.org/10.1007/s10708-014-9599-x>
- Busari, S., & Adebayo, B. (2020). *Nigeria records chloroquine poisoning after trump endorses it for coronavirus treatment*. CNN.com. March 23, 2020 <https://www.cnn.com/2020/03/23/africa/chloroquine-trump-nigeria-intl/index.html>. (Accessed 29 January 2021) accessed April 2, 2020.
- Cao, G., Wang, S., Hwang, M., Padmanabhan, A., Zhang, Z., & Soltani, K. (2015). A scalable framework for spatiotemporal analysis of location-based social media data. *Computers, Environment and Urban Systems*, 51, 70–82. <https://doi.org/10.1016/j.compenvurbsys.2015.01.002>

- Cdc. (2020). COVID data tracker. Retrieved from <https://covid.cdc.gov/covid-data-tracker/#data-tracker-home>. Februray 28, 2021.
- Centers for Disease Control and Prevention. Available at: <https://www.cdc.gov/coronavirus/2019-ncov/> Accessed 4 Februray 2021.
- Chesser, A., Burke, A., Reyes, J., & Rohrberg, T. (2016). Navigating the digital divide: A systematic review of eHealth literacy in underserved populations in the United States. *Informatics for Health and Social Care*, 41(1), 1–19.
- Chou, W. Y. S., Oh, A., & Klein, W. M. (2018). Addressing health-related misinformation on social media. *Jama*, 320(23), 2417–2418.
- Chun, H., Chi, S., & Hwang, B. G. (2017). A spatial disaster assessment model of social resilience based on geographically weighted regression. *Sustainability*, 9(12), 2222.
- Coffee, N. T., Lockwood, T., Rossini, P., Niyonsenga, T., & McGreal, S. (2020). Composition and context drivers of residential property location value as a socioeconomic status measure. *Environment and Planning B: Urban Analytics and City Science*, 47(5), 790–807.
- Crawford, K., & Finn, M. (2015). The limits of crisis data: Analytical and ethical challenges of using social and mobile data to understand disasters. *GeoJournal*, 80(4), 491–502.
- Dom, N. C., Ahmad, A. H., Ishak, A. R., & Ismail, R. (2013). Assessing the risk of dengue fever based on the epidemiological, environmental and entomological variables. *Procedia-Social and Behavioral Sciences*, 105, 183–194.
- Dowd, J. B., Andriano, L., Brazel, D. M., Rotondi, V., Block, P., Ding, X., et al. (2020). Demographic science aids in understanding the spread and fatality rates of COVID-19. *Proceedings of the National Academy of Sciences*, 117(18), 9696–9698. <https://doi.org/10.1073/pnas.2004911117>
- Economist. (2020). *Toplines for March 8–10 economist/YouGov poll*. The Economist. <https://docs.cdn.yougov.com/1ghnpqhpu/econToplines.pdf>. (Accessed 15 April 2020).
- Forman, C., Goldfarb, A., & Greenstein, S. (2018). *How geography shapes-and is shaped by-the internet* (Vol. 269). The new Oxford handbook of economic geography.
- Fotheringham, A. S., Yang, W., & Kang, W. (2017). Multiscale geographically weighted regression (MGWR). *Annals of the Association of American Geographers*, 107(6), 1247–1265. <https://doi.org/10.1080/24694452.2017.1352480>
- Gallardo, R. (2017). *Digital divide index*. Purdue Center for Regional Development.
- Gander, K. (2020). Fauci says social media fuels the spread of disinformation, has impacted pandemic 'more negatively than positively'. Retrieved from <https://www.newsweek.com/fauci-pandemic-social-media-fuels-spread-disinformation-1537856> Accessed Februray 28, 2021.
- Golebiewski, M., & Boyd, D. (2018). *Data voids: Where missing data can easily be exploited*.
- Goodchild, M. F., & Li, L. (2012). Assuring the quality of volunteered geographic information. *Spatial Statistics*, 1, 110–120.
- Greenspan, R. L., & Loftus, E. F. (2021). Pandemics and infodemics: Research on the effects of misinformation on memory. *Human Behavior and Emerging Technologies*, 3(1), 8–12.
- Greenstein, S. (2020). The basic economics of internet infrastructure. *The Journal of Economic Perspectives*, 34(2), 192–214.
- Guess, A., Nagler, J., & Tucker, J. (2019). Less than you think: Prevalence and predictors of fake news dissemination on Facebook. *Science advances*, 5(1), eaau4586.
- Hawkins, B., Kornfield, O'Grady, Copeland, I., & Sonmez, A. (2020). Florida, Texas are latest coronavirus epicenters, the washington post. retrieved from <https://www.washingtonpost.com/nation/2020/06/28/coronavirus-live-updates-us/>.
- Holshue, M. L., DeBolt, C., Lindquist, S., Lofy, K. H., Wiesman, J., Bruce, H., et al. (2020). First case of 2019 novel coronavirus in the United States. *New England Journal of Medicine*. <https://www.nejm.org/doi/full/10.1056/NEJMoa2001191>.
- Imhoff, R., & Lamberty, P. (2020). A bioweapon or a hoax? The link between distinct conspiracy beliefs about the coronavirus disease (COVID-19) outbreak and pandemic behavior. *Social Psychological and Personality Science*, 11(8), 1110–1118.
- Jaeger, P. T., Bertot, J. C., Thompson, K. M., Katz, S. M., & DeCoster, E. J. (2012). The intersection of public policy and public access: Digital divides, digital literacy, digital inclusion, and public libraries. *Public Library Quarterly*, 31(1), 1–20.
- Kasapoglu, A., & Akbal, A. (2020). Relational sociological analysis of uncertainties: The case of COVID-19 in Turkey. *Advances in Social Sciences Research Journal*, 7(4), 197–228. <https://doi.org/10.14738/assrj.74.8116>
- Kouzy, R., Abi Jaoude, J., Kraitem, A., El Alam, M. B., Karam, B., Adib, E., et al. (2020). Coronavirus goes viral: Quantifying the COVID-19 misinformation epidemic on twitter. *Cureus*, 12(3). <https://doi.org/10.7759/cureus.7255>
- Krause, N. M., Freiling, I., Beets, B., & Brossard, D. (2020). Fact-checking as risk communication: The multi-layered risk of misinformation in times of COVID-19. *Journal of Risk Research*, 23(7–8), 1052–1059.
- Kulkarni, P., Prabhu, S., & Ramraj, B. (2020). COVID-19-Infodemic overtaking pandemic? Time to disseminate facts over fear. *Indian Journal of Community Health*, 32(2).
- Lee, J. J., Kang, K. A., Wang, M. P., Zhao, S. Z., Wong, J. Y. H., O'Connor, S., et al. (2020). Associations between COVID-19 misinformation exposure and belief with COVID-19 knowledge and preventive behaviors: Cross-sectional online study. *Journal of Medical Internet Research*, 22(11), Article e22205.
- Lewandowsky, S., Ecker, U. K., & Cook, J. (2017). Beyond misinformation: Understanding and coping with the "post-truth" era. *Journal of applied research in memory and cognition*, 6(4), 353–369.
- Linton, N. M., Kobayashi, T., Yang, Y., Hayashi, K., Akhmetzhanov, A. R., Jung, S.-M., et al. (2020). Incubation period and other epidemiological characteristics of 2019 novel coronavirus infections with right truncation: A statistical analysis of publicly available case data. *Journal of Clinical Medicine*, 9(2). <https://doi.org/10.1101/2020.01.26.20018754>
- Lobato, E. J., Powell, M., Padilla, L. M., & Holbrook, C. (2020). Factors predicting willingness to share COVID-19 misinformation. *Frontiers in Psychology*, 11.
- Marwick, A., & Lewis, R. (2017). *Media manipulation and disinformation online*. New York: Data & Society Research Institute.
- Maxouris, C., Yan, H., & Vera, A. (2021). Coronavirus variant will likely drive a new wave of transmission come spring, some experts say. Retrieved from http://lite.cnn.com/en/article/h_be7903f8a3f04b21928648753a880af3 Accessed Februray 28, 2021.
- Meier, P. (2012). Crisis mapping in action: How open-source software and global volunteer networks are changing the world, one map at a time. *Journal of Map & Geography Libraries*, 8(2), 89–100. <https://doi.org/10.1080/15420353.2012.663739>
- Mollalo, A., Vahedi, B., & Rivera, K. M. (2020). GIS-based spatial modeling of COVID-19 incidence rate in the continental United States. *The Science of the Total Environment*. <https://doi.org/10.1016/j.scitotenv.2020.138884>, 138884.
- Nsoesie, E. O., Oladeji, O., & Sengh, M. D. (2020). Digital platforms and non-communicable diseases in sub-Saharan Africa. *The Lancet Digital Health*, 2(4), e158–e159.
- Office of Rural Health Policy. (2015). *List of rural counties and designated eligible census tracts in metropolitan counties: Updated census 2010*.
- Oh, O., Kwon, K. H., & Rao, H. R. (2010). An exploration of social media in extreme events: Rumor theory and Twitter during the Haiti earthquake 2010. *Iciss*, 231, 7332–7336.
- Oshan, T. M., & Fotheringham, A. S. (2018). A comparison of spatially varying regression coefficient estimates using geographically weighted and spatial-filter-based techniques. *Geographical Analysis*, 50(1), 53–75.
- Oshan, T. M., Li, Z., Kang, W., Wolf, L. J., & Fotheringham, S. (2019). Mgw: A python implementation of multiscale geographically weighted regression for investigating process spatial heterogeneity and scale. *ISPRS International Journal of Geo-Information*, 8(6), 269. <https://doi.org/10.3390/ijgi8060269>
- Oshan, T. M., Smith, J. P., & Fotheringham, A. S. (2020). Targeting the spatial context of obesity determinants via multiscale geographically weighted regression. *International Journal of Health Geographics*, 19, 1–17.
- Ostermann, F. O., & Spinsanti, L. (2011). April). A conceptual workflow for automatically assessing the quality of volunteered geographic information for crisis management. *Proceedings of AGILE*, 2011, 1–6.
- Oxford Analytica. (2020). *Misinformation will undermine coronavirus responses*. Emerald Expert Briefings (oxan-db).
- Pennycook, G., McPhetres, J., Zhang, Y., Lu, J. G., & Rand, D. G. (2020). Fighting COVID-19 misinformation on social media: Experimental evidence for a scalable accuracy-nudge intervention. *Psychological Science*. <https://doi.org/10.1177/0956797620939054>
- Presidential Election Results: Donald J. Trump Wins. (2016). the new York times. Retrieved from <https://www.nytimes.com/elections/2016/results/president>.
- Purwaningsih, T., Prajaningrum, C. S., & Anugrahwati, M. (2018). Building model of flood cases in central java province using geographically weighted regression (GWR). *International Journal of Applied Business and Information Systems*, 2(2), 14–27.
- Rifat, S. A. A., & Liu, W. (2020). Measuring community disaster resilience in the conterminous coastal United States. *ISPRS International Journal of Geo-Information*, 9(8), 469.
- Roesslein, J. (2009). *Tweepy documentation*. Available online <http://docs.tweepy.org/en/v3.5.0/> accessed on February 26th, 2020.
- Roozenbeek, J., Schneider, C. R., Dryhurst, S., Kerr, J., Freeman, A. L., Recchia, G., et al. (2020). Susceptibility to misinformation about COVID-19 around the world. *Royal Society open science*, 7(10), 201199.
- Seo, H., Blomberg, M., Altschwager, D., & Vu, H. T. (2020). *Vulnerable populations and misinformation: A mixed-methods approach to underserved older adults' online information assessment*. New Media & Society, 1461444820925041.
- Seo, H., Erba, J., Altschwager, D., & Geana, M. (2019). Evidence-based digital literacy class for older, low-income African-American adults. *Journal of Applied Communication Research*, 47(2), 130–152.
- Seymour, B., Getman, R., Saraf, A., Zhang, L. H., & Kalenderian, E. (2015). When advocacy obscures accuracy online: Digital pandemics of public health misinformation through an antifluoride case study. *American Journal of Public Health*, 105(3), 517–523.
- Shaw, S. L., Tsou, M. H., & Ye, X. (2016). Human dynamics in the mobile and big data era. *International Journal of Geographical Information Science*, 30(9), 1687–1693. <https://doi.org/10.1080/13658816.2016.1164317>
- Singh, L., Bansal, S., Bode, L., Budak, C., Chi, G., Kawintiranon, K., et al. (2020). *A first look at COVID-19 information and misinformation sharing on Twitter*. arXiv preprint arXiv:2003.13907.
- Stanley, M. L., Barr, N., Peters, K., & Seli, P. (2020). Analytic-thinking predicts hoax beliefs and helping behaviors in response to the COVID-19 pandemic. *Thinking & Reasoning*, 1–14.
- Sui, D., & Goodchild, M. (2011). The convergence of GIS and social media: Challenges for GIScience. *International Journal of Geographical Information Science*, 25(11), 1737–1748. <https://doi.org/10.1080/13658816.2011.604636>
- Sui, D., Goodchild, M., & Elwood, S. (2013). Volunteered geographic information, the exaflood, and the growing digital divide. In *Crowdsourcing geographic knowledge* (pp. 1–12). Dordrecht: Springer. https://doi.org/10.1007/978-94-007-4587-2_1.
- Székel, G. J., & Rizzo, M. L. (2014). Partial distance correlation with methods for dissimilarities. *Annals of Statistics*, 42(6), 2382–2412, 10.1214/14-AOS1255 <https://projecteuclid.org/euclid.aos/1413810731>.
- Székel, G. J., & Rizzo, M. L. (2016). Partial distance correlation. In R. Cao, W. González Manteiga, & J. Romo (Eds.), *Nonparametric statistics. Springer proceedings in mathematics & statistics* (Vol. 175). Cham: Springer. https://doi.org/10.1007/978-3-319-41582-6_13.

- Székely, G. J., Rizzo, M. L., & Bakirov, N. K. (2007). Measuring and testing dependence by correlation of distances. *Annals of Statistics*, 35(6), 2769–2794. [10.1214/009053607000000505](https://projecteuclid.org/euclid.aos/1201012979) <https://projecteuclid.org/euclid.aos/1201012979>.
- Timmis, K., & Brüssow, H. (2020). The COVID-19 pandemic: Some lessons learned about crisis preparedness and management, and the need for international benchmarking to reduce deficits. *Environmental Microbiology*. <https://doi.org/10.1111/1462-2920.15029>
- Törnberg, P. (2018). Echo chambers and viral misinformation: Modeling fake news as complex contagion. *PloS One*, 13(9), Article e0203958.
- Tsou, M. H. (2015). Research challenges and opportunities in mapping social media and Big Data. *Cartography and Geographic Information Science*, 42(sup1), 70–74. <https://doi.org/10.1080/15230406.2015.1059251>
- Uscinski, J. E., Enders, A. M., Kloststad, C., Seelig, M., Funchion, J., Everett, C., et al. (2020). Why do people believe COVID-19 conspiracy theories? *Harvard Kennedy School Misinformation Review*, 1(3).
- Vosoughi, S., Roy, D., & Aral, S. (2018). The spread of true and false news online. *Science*, 359(6380), 1146–1151. <https://doi.org/10.1126/science.aap9559>
- Waldrop, T., Alsup, D., & McLaughlin, E. (2020). *Fearing coronavirus, Arizona man dies after taking a form of chloroquine used to treat aquariums*. CNN.com. March 25, 2020 <https://www.cnn.com/2020/03/23/health/arizona-coronavirus-chloroquine-death/index.html>. (Accessed 29 January 2021).
- Wang, G., Minnis, R. B., Belant, J. L., & Wax, C. L. (2010). Dry weather induces outbreaks of human West Nile virus infections. *BMC Infectious Diseases*, 10(1), 1–7.
- Watkins, R. E., Eagleson, S., Beckett, S., Garner, G., Veenendaal, B., Wright, G., et al. (2007). Using GIS to create synthetic disease outbreaks. *BMC Medical Informatics and Decision Making*, 7(1), 1–14. <https://doi.org/10.1186/1472-6947-7-4>
- Whitacre, B. E., & Mills, B. F. (2007). Infrastructure and the rural—urban divide in high-speed residential Internet access. *International Regional Science Review*, 30(3), 249–273.
- Xiao, Y., Huang, Q., & Wu, K. (2015). Understanding social media data for disaster management. *Natural Hazards*, 79(3), 1663–1679. <https://doi.org/10.1007/s11069-015-1918-0>
- Xie, M., & Chen, Q. (2020). Insight into 2019 novel coronavirus — an updated interim review and lessons from SARS-CoV and MERS-CoV. *International Journal of Infectious Diseases*, 94, 119–124. <https://doi.org/10.1016/j.ijid.2020.03.071>
- Yang, K. C., Torres-Lugo, C., & Menczer, F. (2020). *Prevalence of low-credibility information on twitter during the covid-19 outbreak*. arXiv preprint arXiv:2004.14484.
- Ziemke, J. (2012). Crisis mapping: The construction of a new interdisciplinary field? *Journal of Map & Geography Libraries*, 8(2), 101–117. <https://doi.org/10.1080/15420353.2012.662471>
- Zook, M., Graham, M., Shelton, T., & Gorman, S. (2010). Volunteered geographic information and crowdsourcing disaster relief: A case study of the Haitian earthquake. *World Medical & Health Policy*, 2(2), 7–33. <https://doi.org/10.2202/1948-4682.1069>