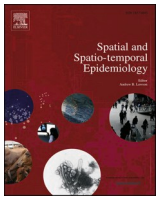
Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Spatial and Spatio-temporal Epidemiology 42 (2022) 100521

Contents lists available at ScienceDirect

Spatial and Spatio-temporal Epidemiology

journal homepage: www.elsevier.com/locate/sste



Space-time analysis of COVID-19 cases and SARS-CoV-2 wastewater loading: A geodemographic perspective

J.R. Nelson [a,*], A. Lu [b], J.P. Maestre [b], E.J. Palmer [b], D. Jarma [b], K.A. Kinney [b], T.H. Grubesic [c], M.J. Kirisits [b]

[a] Department of Geosciences, Auburn University, 2050 Beard Eaves Coliseum, Auburn, AL 36849, USA
[b] Department of Civil, Architectural, and Environmental Engineering, The University of Texas at Austin, 301 E. Dean Keeton St., Stop C1786, Austin, TX 78712, USA
[c] Geoinformatics & Policy Analytics Laboratory, School of Information, University of Texas at Austin, USA

A B S T R A C T

Severe acute respiratory syndrome - coronavirus 2 (SARS-CoV-2) continues to effect communities across the world. One way to combat these effects is to enhance our collective ability to remotely monitor community spread. Monitoring SARS-CoV-2 in wastewater is one approach that enables researchers to estimate the total number of infected people in a region; however, estimates are often made at the sewershed level which may mask the geographic nuance required for targeted interdiction efforts. In this work, we utilize an apportioning method to compare the spatial and temporal trends of daily case count with the temporal pattern of viral load in the wastewater at smaller units of analysis within Austin, TX. We find different lag-times between wastewater loading and case reports. Daily case reports for some locations follow the temporal trend of viral load more closely than others. These findings are then compared to socio-demographic characteristics across the study area.

Introduction

It has been over two years since the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), which causes coronavirus disease 2019 (COVID-19), caused a worldwide lockdown (WHO, 2020). Unlike previous outbreaks of human pathogenic viruses (e.g., SARS-CoV-1 in 2003, H1N1 in 2009, and the Middle East respiratory syndrome [MERS] coronavirus epidemic between 2012 and 2015), the high transmissivity rate of SARS-CoV-2 made this outbreak unique (Fauci et al., 2020; Godri Pollitt et al., 2020; Petersen et al., 2020; Rodpothong and Auewarakul, 2012). At least part of the high transmissivity is attributable to the lag between viral shedding and the onset of physical symptoms (Petersen et al., 2020). This lag, coupled with the substantial hurdles associated with large-scale testing early in the pandemic, made quarantine measures less effective in reducing transmission (Kucharski et al., 2020). Although many countries have implemented (and continue to implement) physical distancing requirements, mask mandates, and restrictions on indoor and outdoor gatherings, more than 532 million COVID-19 cases have been confirmed worldwide (JHU, 2021). However, many scientists believe the reported cases underestimate the actual count due to problems with testing (availability, false negatives), asymptomatic individuals, and a reluctance for testing by some symptomatic individuals (Alwan, 2020; Tanne, 2020; Wu et al., 2020). Alwan (2020) suggests that one of the main factors contributing to testing reluctance is the financial impact of missing work for those without paid time off or sick leave.

The ability to work remotely influences the financial impact of a positive COVID-19 test. For people who cannot work remotely after a positive COVID-19 test, the financial impacts could be devastating, while those who can work remotely likely feel much less of an impact. Remote work is not an option for many industries focusing on service (e.g., restaurants, hotels, janitorial services), construction, or the gig-economy (e.g., Instacart, Uber), among many others. A positive COVID-19 test for such a worker could mean that they go without pay. For most people, losing income is not a viable option. This potential loss means that many individuals choose to work, regardless of COVID-19 status. Compounding these challenges is that employees in these aforementioned industries skew towards lower-income minority groups (Goldman et al., 2021). As a result, recent research finds a disproportionate representation of minority populations in COVID-19 cases, hospitalizations, and deaths (S. J. Kim and Bostwick, 2020; Muñoz-Price et al., 2020; Wadhera et al., 2020).

Due to testing difficulties (e.g., availability), reluctance to test, asymptomatic individuals, and the lag between viral shedding and physical symptoms, it is challenging to accurately identify COVID-19 community case counts (Wu et al., 2020). This uncertainty has made mounting an appropriate response strategy problematic (Noh and Danuser, 2021). To that end, a growing number of researchers are

* Corresponding author.

https://doi.org/10.1016/j.sste.2022.100521
Received 26 October 2021; Received in revised form 24 May 2022; Accepted 26 May 2022
Available online 28 May 2022

turning to wastewater epidemiology to assess viral loads in a sewershed (Larsen and Wigginton, 2020). This work has primarily focused on developing surveillance tools to test for SARS-CoV-2 RNA (Ahmed et al., 2020; La Rosa et al., 2020), identify ideal sampling points within a sewer network (Balboa et al., 2021; Yeager et al., 2021), and compare SARS-CoV-2 concentrations in wastewater with reported cases (Medema et al., 2020; Randazzo et al., 2020). Results from these studies highlight the value of wastewater in monitoring the prevalence of COVID-19 within a sewershed and the assocaited communities. However, from a response and mitigation perspective, the size of a sewershed makes it difficult to pinpoint specific sub-communities for treatment, testing, and supplies without supplementing sampling at the wastewater treatment plant with sampling of key manholes in the sewershed.

Increasingly, studies are exploring the efficacy of wastewater monitoring to provide actionable insights for COVID-19 response. This includes the recent work by Scott et al. (2021) who utilized gene markers to monitor the spread of COVID-19 on a U.S. college campus, Pillay et al. (2021) who provided a unique perspective on wastewater monitoring in South Africa, and Saththasivam et al. (2021) who were the first to measure SARS-CoV-2 RNA fragments in wastewater in Qatar. This work continues to evolve with some of the most recent studies taking place internationally, such as in Madrid (Lastra et al., 2022). For more information on the state of the art in monitoring we recommend the recent review by Lahrich et al. (2021). In this study, we aim to contribute to this growing body of research by exploring the spatial variation in SARS-CoV-2 loading at wastewater treatment plants and daily reported COVID-19 cases at the ZIP code level in Austin, Texas. To do this, we begin by exploring the spatio-temporal variation of daily COVID-19 cases for each ZIP code falling within two sewershed catchment areas. We then compare the temporal relationship of SARS-CoV-2 wastewater loading (N-gene copies/day) and daily reported COVID-19 cases to uncover the temporal dynamics of viral shedding and case reporting at a more granular spatial scale. Our cross-correlation time-series analysis identifies the presence of lead periods between spikes in SARS-CoV-2 loading and reported cases when considering the aggregate number of cases in the sewershed and wastewater loadings. However, when we disaggregate and apportion cases to the individual ZIP codes based on the population within the sewershed, the lag periods are generally the same, but variation in the strength of correlations between ZIP code case reports and loading levels become clear. The results of this paper support the use of wastewater analysis as a surveillance tool and highlight how case disaggregation and geospatial analysis can enhance surveillance.

Background

The concept of community vulnerability connects to anthropogenically induced or natural hazards – typically large pollution releases, hurricanes, tornados, and other events with the potential to cause significant damage to the built and natural environment. In this context, vulnerability is defined as a community's susceptibility to harm or the potential for loss within the community (Cutter, 1996; Weichselgartner, 2001). Building a deeper understanding of the link between community vulnerability and hazard risk is critical for minimizing the damage caused by disasters – whether anthropogenic or natural.

The ongoing COVID-19 pandemic, by any standard, is a disaster. Millions of people have died (JHU, 2021), the pandemic continues to cripple economies (Bauer et al., 2020), and the ripple effects of the pandemic will manifest for many years (British Academy, 2021; Horowitz et al., 2021). Simply put, the world was woefully underprepared for an event of this magnitude, as evidenced by a general reluctance to act on early information about the virus (Caduff, 2020), the rampant spread of misinformation (H. K. Kim et al., 2020), lack of testing (Caduff, 2020), scarcity of medical supplies (Ranney et al., 2020), and general supply-chain problems (Guan et al., 2020). Moreover, the virus exposed many inadequacies associated with our collective ability to respond to this type of event, highlighting how the most vulnerable communities face the highest risk of COVID-19 illness and death, due, in part, to difficulties associated with monitoring community spread and prevalence (S. J. Kim and Bostwick, 2020).

Socioeconomic health inequities of COVID-19

Post-hoc research on COVID-19 case rates reveals patterns of effects between COVID-19 and socio-demographics that parallel similar research on other environmental and social inequities. The burdens fueled by COVID-19 have fallen disproportionately on vulnerable and marginalized communities. This finding appears to be scale agnostic, with research confirming these patterns across cities (DiMaggio et al., 2020), counties (Liao and De Maio, 2021), states (Karaca-Mandic et al., 2021), and regions (Strully et al., 2021). These burdens range from higher percentages of positive tests to increased hospitalizations and mortality.

More specifically, recent work in Louisiana found that Black Americans made up 70.6% of the COVID-19 hospitalizations and 76.3% of COVID-19 deaths, even though only 36% of the population for the study area is Black (Price-Haywood et al., 2020). Likewise, in Milwaukee, Wisconsin, a cross-sectional analysis revealed that Black males were at higher risk of testing positive for COVID-19 and being hospitalized due to COVID-19 complications than were their racial and ethnic counterparts (Muñoz-Price et al., 2020). Almagro & Orane-Hutchinson (2020) reported similar results, identifying a significant, positive relationship between low-income majority-minority areas and the percentage of positive COVID-19 tests in New York City. Furthermore, Mahajan & Larkins-Pettigrew (2020) and Karmakar et al. (2021) both show the existence of significant racial inequities in COVID-19 mortality using data from across the United States.

Many medical professionals point to the health inequities in the United States as the underlying reason for some of the racial disparities in COVID-19-related health outcomes (Bibbins-Domingo, 2020). For example, pre-existing conditions can make the risk of COVID-19 illness greater. Because minority communities are disproportionately affected by diabetes, hypertension, obesity, and other ailments (Towne et al., 2017; Zarefsky, 2020), their mortality risk for COVID-19 is much higher. In addition to comorbidities, members of minority groups are more likely to be part of the low-wage essential workforce, such as bus drivers, custodians, service workers, and other front-line industry workers (Bibbins-Domingo, 2020). Indeed, Almagro & Orane-Hutchinson (2020) tested and confirmed that working in transportation, construction, and other service industry occupations was positively and significantly associated with positive COVID-19 tests. Because these workers typically do not have paid time off or sick leave (Schneider and Harknett, 2021), they might feel less inclined to be tested or stay home from work following a positive test or symptom onset. As a result, the actual COVID-19 case count might be much higher in such communities. Thus, even though free COVID-19 testing is available, testing is unlikely to be the most reliable means for monitoring COVID-19 prevalence within a community.

Finally, recent research suggests that commercial activity, specifically the movement of people between work and home, is an important explanatory factor in the spread of COVID-19 (Bontempi et al., 2021;

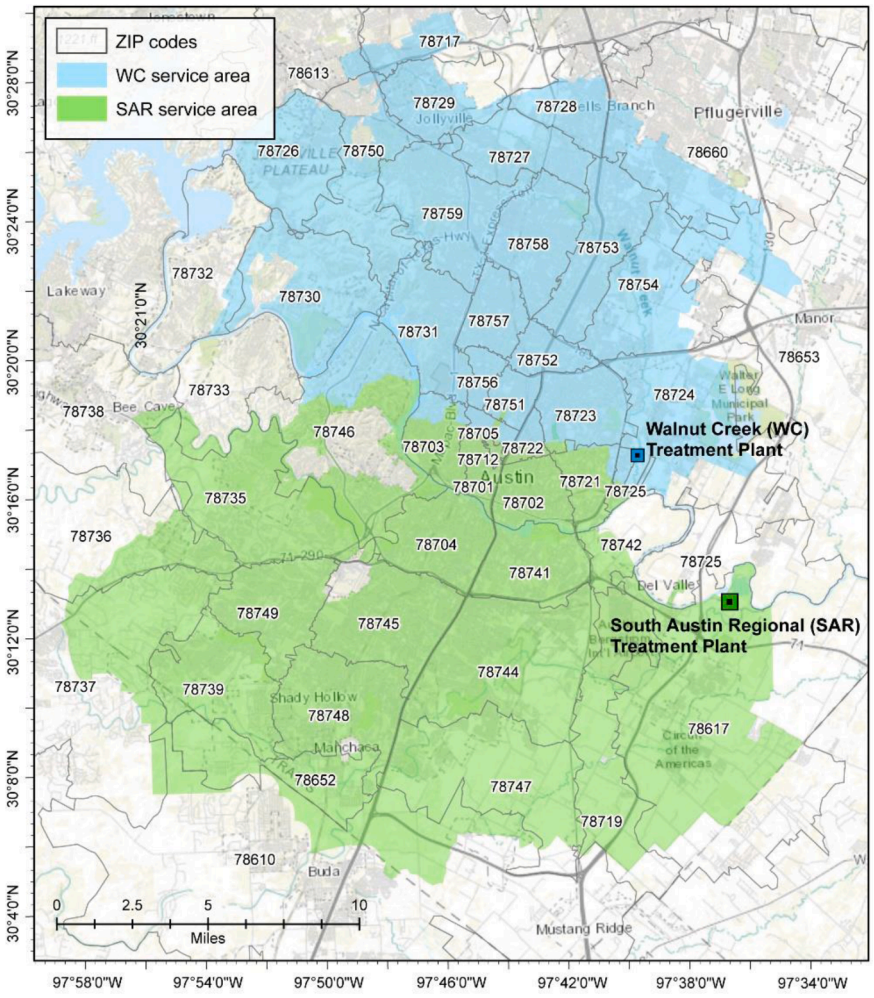


Fig. 1. Austin study area with the two wastewater treatment service areas highlighted.

Grubesic et al., 2021). Indeed, recent studies have demonstrated an interesting link between COVID-19 transmission in communities where individuals are likely to work at "risky" facilities with the potential for high infection rates (Towers et al., 2022). In short, individuals who commute to and from jobs that place them at higher risk of infection are in-essence two-way vectors for transmission. They may transport the virus to and from the locations where they work and live.

Proactive monitoring

Proactive monitoring is essential for preparedness and effective resource allocation (Lin Moe and Pathranarakul, 2006). Due to the concern regarding underestimation of the actual case count and level of community spread of COVID-19, scholars have been advancing SARS-CoV-2 monitoring in wastewater. Farkas et al. (2020) detailed that viral concentrations in wastewater can provide an essential indicator of viral disease prevalence in a community. While the methods for concentrating and detecting SARS-CoV-2 in wastewater continue to be developed, many researchers have successfully extracted SARS-CoV-2 RNA from untreated sludge samples or other bulk wastewater sampling points (Balboa et al., 2021; Palmer et al., 2021). Quantifying the viral concentration in wastewater samples typically involves reverse-transcription, quantitative, real-time polymerase chain reaction (RT-qPCR), digital PCR, or digital-drop PCR (ddPCR) methods. For example, Ahmed et al. (2020) employed RT-qPCR to determine the SARS-CoV-2 concentration in wastewater to estimate the prevalence of COVID-19 among those who lived within the target Australian sewershed, while (Liu et al., 2020) reported that the ddPCR test returned fewer false negatives in low viral load specimens than did RT-qPCR methods. The results from Suo et al. (2020) also suggest that ddPCR might be better than RT-qPCR for minimizing false negatives.

Wastewater-based epidemiology lends itself to an early-warning and preparedness system for decision-makers and health officials. For example, several recent studies document the presence of SARS-CoV-2 in wastewater *weeks* before the first case was actually reported (La Rosa et al., 2020; Medema et al., 2020; Randazzo et al., 2020). Moreover, results from other studies suggest that one can use wastewater monitoring to track the prevalence of the virus as it ebbs and flows within a community over time. For example, Ahmed et al. (2020) explored this temporal dynamic and found some initial alignment between the decline in the first COVID-19 case wave and wastewater loading. Kumar et al. (2021) also published promising results, graphically depicting a correlation between wastewater SARS-CoV-2 concentrations and reported COVID-19 cases with a wastewater lead time of 1–2 weeks. Weidhaas et al. (2021) report a similar wastewater lead time.

Interestingly, Nemudryi et al. (2020) documented an even more dynamic relationship that varied depending on a waning (decreasing) or waxing (increasing) of COVID-19 case numbers. During the waning period (decrease in cases after initial surge), patient-reported symptom onset preceded *detectable* levels of SARS-CoV-2 in the wastewater by

eight days,[3] with a spike in positive COVID-19 tests two days after wastewater detection. Conversely, during a waxing phase (resurgence), symptom onset was found to precede wastewater detection by five days, with positive tests following four days after SARS-CoV-2 was detected in the wastewater. Given that this analysis was performed in the relatively small city of Bozeman, Montana, the lag between symptom onset and wastewater detection might be longer than expected due to the difficulty in detecting SARS-CoV-2 in wastewater for the small COVID-19-positive population.

These studies highlight the potential for wastewater monitoring as an early-warning system and method for estimating the prevalence of COVID-19 within the communities served by a given sewershed. However, one drawback of current approaches is the lack of explicitly accounting for the number of people or facilities that are contributing to the wastewater. For example, consider the equation to estimate viral gene copies used in this analysis:

$$\text{CoV} - 2 \text{ loading} = \left(\frac{gene\ copies}{L\ wastewater}\right) * \left(\frac{L\ wastewater}{day}\right) \quad (1)$$

Eq. (1) can then be transformed to estimate the number of ill individuals within the sewershed (Weidhaas et al., 2021) with some assumptions about how much effluent each individual contributes. That said, the estimate of total number of infected individuals based on SARS-CoV-2 concentration critically depends on knowing the size of the population contributing wastewater in the sewershed.

Unfortunately, sewersheds rarely align with administrative boundaries; instead, they often bisect locations where the population is known (e.g., census tracts, counties, ZIP codes). This is because many sewer systems are designed for gravity wastewater collection systems (GWCS) or pressure collection. The former relies heavily on the elevation changes across the landscape to control where and how wastewater flows through the system (Islam, 2017). The latter depends on a series of pumps to move the water through the system. To make the system as efficient as possible, they are designed and modified to optimize flow, which will not always conform to administrative boundaries (Miszta-Kruk, 2016). There also are reporting challenges with COVID-19 cases. Agencies typically assign COVID-19 case counts to each administrative unit (e.g., tract or ZIP code). As a result, researchers interested in estimating case counts or positive case rates based on gene copies of SARS-CoV-2 in wastewater must apportion the population and the number of cases to match the service area of the sewershed. If one has information about the population with sufficient granularity, it is possible to geocode (i.e., add latitude and longitude coordinates) household addresses to provide a highly accurate estimate of the population within the sewershed (Weidhaas et al., 2021). However, this level of granularity is not always available, leading to the use of rough population estimates based on large regional areas (Ahmed et al., 2020).

While wastewater-based epidemiology provides an initial indication of COVID-19 community prevalence, it remains challenging to identify and effectively direct intervention resources to sub-communities that might be more vulnerable to COVID-19 and actively experiencing community spread. These challenges are especially acute when the sewersheds bisect administrative units, which is the case for Austin, Texas (Fig. 1). This work explores an approach that will help facilitate a more spatially granular assessment of COVID-19 cases and SARS-CoV-2 loading in wastewater. To accomplish this task, we utilize daily reported COVID-19 cases in Austin, Texas, and SARS-CoV-2 wastewater data from the two major sewersheds in Austin (Fig. 1). We begin by calculating space-time clusters of COVID-19 case prevalence across the ZIP codes in Austin, TX. Then, we perform a cross-correlation analysis between SARS-CoV-2 loading in wastewater and daily reported COVID-19 cases for the proportion of each ZIP code that falls within the sewershed of interest. Not only does this allow us to determine the lead time between viral shedding and reported cases, but it also allows us to determine whether the reported cases in some ZIP codes follow the wastewater loadings more closely than others. If so, it would provide a more nuanced understanding of which communities are experiencing more extensive community spread and require more response resources.

Data and methods

Data collection

We obtained case information from the City of Austin's COVID-19 dashboard (APH, 2021). Reported cases are laboratory-confirmed, official city numbers and are reported at the ZIP code level. We gathered case data from May 2020 through January 2021 for each ZIP code in Austin. We transformed the reported cumulative cases to reflect the new daily cases by subtracting the previous day's cumulative case count (t-1) from the current day's cumulative case count (t).

Wastewater collection

Flow-weighted, 24-hour composite samples of primary clarifier effluent (500–1000 mL) were collected from May 2020 – January 2021 from the two major wastewater treatment plants in Austin, Texas: Walnut Creek (WC) and South Austin Regional (SAR) wastewater treatment plants (Fig. 1). The WC sewershed covers 169.2 sq. miles and processed approximately 53 million gallons per day (MGD) during the study period. The SAR sewershed covers 251.2 sq. miles and processed approximately 43 MGD during the study period (Palmer et al., 2021). Samples generally were collected three times per week at the WC wastewater treatment plant and two times per week at the SAR wastewater treatment plant. Samples were transported on ice to the laboratory at the University of Texas at Austin, mixed, divided into 50-mL aliquots, and stored at −20 °C.

Sample processing and RNA extraction

The wastewater samples were processed and RNA extracted according to the recommended protocol in Palmer et al. (2021). Briefly, triplicate 50-mL aliquots for each wastewater sample were placed in a room temperature water bath for 1.5 h to thaw. Aliquots were then pasteurized at 60 °C for 1.5 h. Two-phase centrifugation at 4500 xg for 5 min at 4 °C resulted in a pellet that was subsequently resuspended in the lysis buffer provided in the MagMax Microbiome Ultra Nucleic Acid Isolation kit (ThermoFisher, Waltham, Massachusetts). Homogenization was conducted in the lysing tubes of the kit via four rounds of disruption in a FastPrep-24 (MP Biomedicals; Santa Ana, California) at 4.0 m/s for 20 s and centrifugation at 13,000 xg for 15–20 s.

RNA was extracted with the MagMax Microbiome Ultra Nucleic Acid Isolation kit according to the manufacturer's instructions utilizing the KingFisher Flex system (ThermoFisher, Waltham, Massachusetts). Extracted RNA was stored at −20 °C. The nucleic acid concentration was measured spectrophotometrically with the Synergy Neo 2 Hybrid Multi-Mode Reader (BioTek; Winooski, Vermont).

SARS-CoV-2 quantification by RT-qPCR

The SARS-CoV-2 concentrations of aliquots were determined via

[3] All laboratory confirmed cases of COVID-19 were contacted via telephone and asked a series of questions, including the date they first started to feel symptoms.

triplicate RT-qPCR reactions on each RNA extract, using a ViiA7 Real-Time PCR System (ThermoFisher; Waltham, Massachusetts) and the CDC nCOV_N2 primer/probe set (Integrated DNA Technologies [IDT], Coralville, Iowa). A standard curve (20,000, 2000, 200, 20, and 2 N gene copies/μL) was prepared via a serial dilution of the 2019-nCoV_N_Positive Control (IDT). RT-qPCR reactions had a total volume of 20 μL, and contained: 5 μL of TaqMan™ Fast Virus 1-Step Master Mix (ThermoFisher; Waltham, Massachusetts), 1.49 μL of the CDC nCOV_N2 primer/probe set mix (for a final concentration of 500 nM of each primer and 125 nM of probe), 8.51 μL of PCR-grade water, and 5 μL of RNA extract. The thermal cycler conditions recommended by the CDC were used: 50 °C for 5 min, 95 °C for 20 s, and 40 cycles of 95 °C for 15 s and 60 °C for 60 s. The limit of detection (LOD) of triplicate N2 assays was determined to be 7.5 N gene copies per reaction via six serial dilutions (100,000 to 1 copy per reaction) of the 2019-nCoV_N_Positive Control SARS-CoV-2 plasmid standard and subsequent 20 replicates of standards with 10, 5, 2.5, and 1 copy per reaction. Because we utilized 50-mL sample aliquots and 200-μL elution buffer, the LOD translates to 6000 N gene copies/L of wastewater. RT-qPCR reactions that did not yield amplification were not assigned a concentration value, and the lack of amplification was noted per the protocol followed by Ahmed et al. (2020a). Samples that did not yield an average concentration of the N gene above the LOD were not included in the correlational analyses. The validity of RT-qPCR results was ensured by including positive controls (2019-nCoV_N_Positive Control, IDT) and negative controls (PCR-grade water) in every RT-qPCR plate. The SARS-CoV-2 concentration in each extract was calculated in units of N gene copies/μL extract by averaging the triplicate reactions for each extract. The average SARS-CoV-2 concentration of each wastewater sample in units of N gene copies/L wastewater was calculated by using the aliquot volume (50 mL), the volume of elution buffer (200 μL), and the concentrations of the extracts. Finally, the average concentration of SARS-CoV-2 is calculated for each wastewater sample and multiplied by the plant's wastewater flowrate for that day (L/d) to obtain the loading of SARS-CoV-2 (N gene copies/d) (Appendix A).

Case data preparation

The area served by each treatment plant and each ZIP code in the Austin area is shown in Fig. 1. It is important to note that several ZIP codes on the periphery of Austin only partially overlap with the sewersheds. Therefore, an essential first step in analyzing the correspondence between SARS-CoV-2 wastewater loading and COVID-19 case count is to apportion cases to each ZIP code based on the percent of the population in each ZIP code that falls within the sewershed area. Several alternative approaches can facilitate this apportionment. First, one could take a purely area-based approach. This approach scales cases by the percent area of each ZIP code that overlaps with the sewershed. However, this simplistic approach fails to account for locations with sparsely (or densely) populated areas. Second, one could use the population of nested administrative units (e.g., block groups or tracts within a ZIP code) that fall within the sewershed to apportion cases. This alternative is more accurate than the area-based approach but will likely face similar issues as the first approach if the sewershed bisects the smaller administrative units, if the administrative unit is sparsely populated, or if the population concentrates in a small area of a larger administrative unit. Third, one could utilize the locations of structures in the administrative unit. Here, the cases are scaled to reflect the proportion of structures within each administrative unit for the sewershed. This approach provides a more accurate representation of population locations while also directly accounting for wastewater sources.

We employed the third approach and began by clipping each ZIP code area to match the spatial bounds of the sewershed. Next, we used a comprehensive structure shapefile to identify the count of structures within each ZIP code and the sewershed boundaries. Next, we divided the number of structures falling within the sewershed portion of the ZIP code (S_{struct}) by the total number of structures in the ZIP code (S_{total}) to build a scaling factor for case counts. Specifically, we multiplied this proportion by the total daily case count reported for each ZIP code (2), then, after scaling the cases, we calculated the daily case rate for each ZIP code (or ZIP code portion overlapping the sewershed area) at time (t) for each ZIP code (i) per 10,000 residents (3).

$$apportioned\_cases = \frac{S_{struct}}{S_{total}} * daily\ case\ count \tag{2}$$

$$COVID-19\ Case\ Rate_i^t = \left(\frac{daily\ cases_i^t}{population_i}\right) * 10,000 \tag{3}$$

The case rate (3) for each time period (t) for each ZIP code (i) was used for the space-time cluster analysis to facilitate a more objective comparison of COVID-19 cases across areas with different populations.

Daily cases – wastewater correlation analysis

Sewershed level

We performed cross-correlations between the wastewater loading at each treatment plant (N gene copies/day) and the apportioned new daily cases[4] in aggregate (total for all ZIP codes in the sewershed area) and to each ZIP code individually. Where the wastewater data are concerned, the WC wastewater treatment plant contained three dates (10/13/20, 10/14/20, 10/15/20) with two loading values. For these cases, we took an average of the two values. In addition, loading data from 11/23/20 – 12/2/20 and 12/9/20 – 12/30/20) were omitted from both wastewater datasets because the samples collected on these days had a holding time at 4 °C that exceeded 7 days. We used loading and case data collected between 7/1/20 and 1/11/21 for the SAR sewershed and 5/17/20 to 1/6/21 for WC sewershed.

For the sewershed area analysis, we plotted SARS-CoV-2 loading data for WC and SAR against new daily reported cases. We applied a temporal lag ($t+1$, $t+2$,…$t+10$) to the reported date for the loading data in order to shift it forward in time because SARS-CoV-2 detection in wastewater often precedes symptom onset and case reporting. Next, we used case data corresponding to the date of the lagged loading values and calcualted a Pearson correlation to compare. For example, when analyzing a lag of one day, the COVID-19 case data from 7/2/20 would be paired with the 7/1/20 loading data.

ZIP code analysis

We began the ZIP code level analysis by z-score standardizing both the COVID-19 case counts and SARS-CoV-2 wastewater loadings. We then fit smoothing curves to both datasets using the locally weighted scatterplot smoothing (LOWESS) technique (Nemudryi et al., 2020). The model uses a local weighting value to determine the influence that surrounding points have when determining the value of the smoothed estimate. After several iterations utilizing different weights, we determined that 0.20 offered a good balance between over and underfitting for the wastewater data and 0.166 for the case data. Next, the wastewater data was iteratively lagged relative to the case data in one-day increments beginning at t-15 days and ending and $t+15$ days. Next, we calculated a Pearson cross-correlation value between wastewater loading and daily case numbers for each lag period, subsequently recording the lag at which the highest cross-correlation value (max synchrony) was found. Finally, we added both of these measures as an attribute of the ZIP code shapefile for mapping. The lag at which max

[4] Hereafter any reference to cases is related to the apportioned numbers.

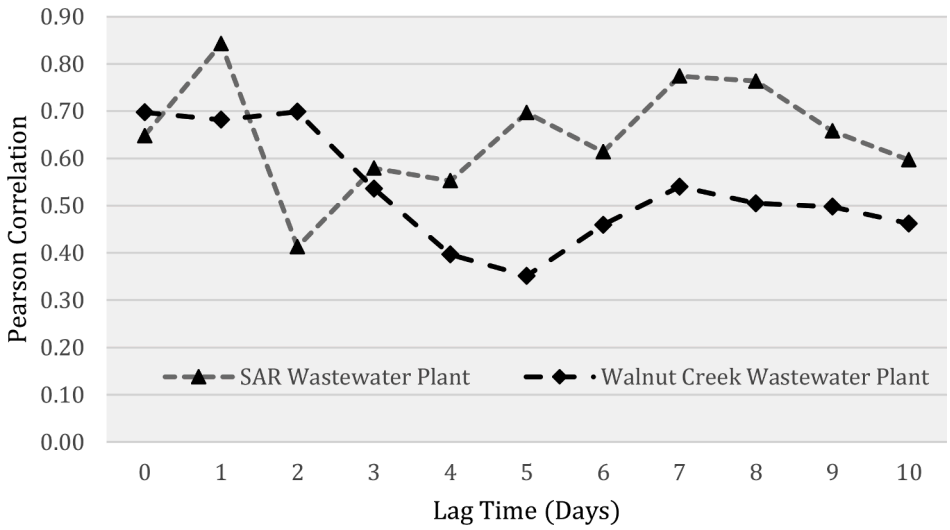


Fig. 2. Correlation between SARS-CoV-2 wastewater loading and apportioned aggregate case count for each sewershed at different time lags. Lag time indicates how far wastewater data precedes case data.

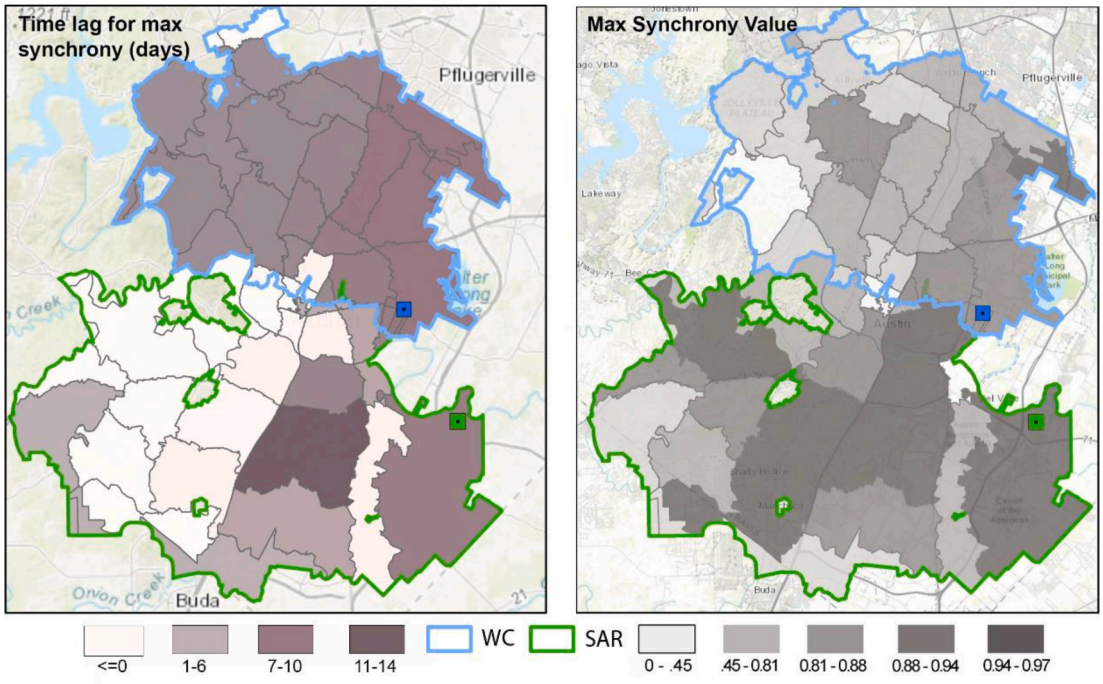


Fig. 3. The max synchrony value represents the highest correlation between wastewater loading and case count for each ZIP code. The reported time lag indicates the day (lag) associated with the highest correlation between wastewater loading values and case data. Sewersheds are outlined in green and blue.

synchrony occurs provides a general measure of how far ahead wastewater loading preceded case reporting, while the max synchrony value itself will measure how strong the relationship between wastewater loading and reported cases is for each ZIP code.

Space-time analysis

We also performed a space-time cluster analysis of the daily COVID-19 case rates using apportioned case numbers to identify the spatial and temporal clusters of COVID-19 across the sewershed ZIP codes. We employed a space-time variation of the local Moran's *I* statistic, which determines the percentage of time that each ZIP code was a member of a hotspot (high case numbers surrounded by areas of similarly high case numbers), cold spot (low case numbers surrounded by areas of similarly low case numbers), or outlier (low cases surrounded by high cases, or vice versa) throughout our study period. The local Moran's *I* is defined as follows:

$$I_i = \frac{x_i - \overline{X}}{S_x^2} \sum_{j=1, j\neq i}^{n-1} w_{i,j}\left(x_j - \overline{X}\right) \tag{4}$$

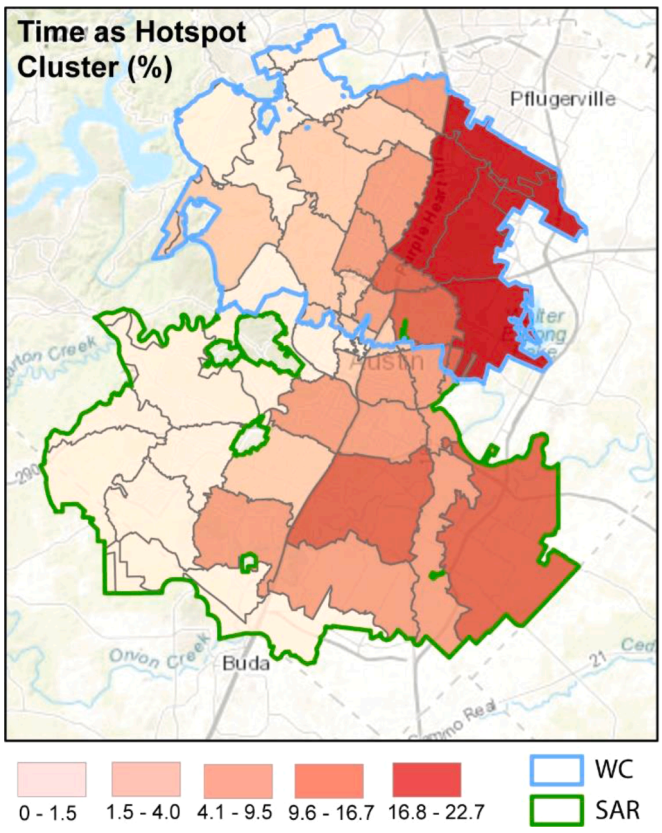


Fig. 4. Graphical display of the spatio-temporal COVID-19 hotspot clusters for each ZIP code. Coloring corresponds to the amount of time that each ZIP code was a members of a space-time case hotspot denoted as a percent of the total days (320) of reported case data used in the analysis.

Where x_i is the attribute for feature i, $\overline{X}$ is the global average of attribute x, w_{ij} is the weight between feature i and j, and S_x^2 is the variance for attribute x. We classify a feature (i.e., ZIP code area) as part of a hotspot, cold spot, outlier, or not significant based on the value of I and its associated z-score and p-value. In addition to spatial neighbors j, the space-time variation of the local Moran's I considers temporal neighbors (z) at a lag of two-time steps, allowing us to evaluate the persistence of clusters across both space and time. We evaluate these results in light of several socioeconomic measures, including a social vulnerability index calculated by the Centers for Disease Control (CDC).[5]

Results

Once we apportioned the new daily cases to each ZIP code, we conducted cross-correlations between cases and wastewater loading at the aggregate sewershed level and at the ZIP code level. Both of these were performed according to the respective sewershed. When using case data aggregated to the sewershed level, the cross-correlations revealed a significant positive correlation between loading and new daily cases for both sewersheds (WC: $r = 0.697$, p-value <0.001; SAR: $r = 0.648$, p-value <0.001). Fig. 2 displays the lag time plot and correlation coefficients. The results indicate a significant positive correlation (p-value < 0.05) for all lag times (Appendix B); however, the correlation coefficient was the highest when wastewater was shifted backwards in time by 1–2-days for WC and 1-day for SAR. Both treatment plants had a second local maxima in correlation coefficient when wastewater data preceded cases by 7- and 8-days (Fig. 2). These maxima align with previous findings, which reported SARS-CoV-2 trends in wastewater appearing 2–8 days before reported COVID-19 cases (Karthikeyan et al., 2021).

We then performed a similar cross-correlation analysis but used the apportioned case data for each ZIP code. Using the ZIP code as the unit of analysis allowed us to map the correlation value for each ZIP code and the lag when max synchrony occurred (Fig. 3). The results reveal variation in the strength of the correlation and the lag when the correlation is strongest. Overall, the correlation between wastewater loading and new daily reported cases for each ZIP code was generally high, with an average Pearson value of 0.79 ($p < 0.05$) and 0.85 ($p<0.05$) for WC and SAR, respectively.

Both sewersheds share a certain degree of heterogeneity in max synchrony values (right panel Fig. 3). There was a more noticeable east to west trend associated with the max synchrony values in the WC compared to SAR, with the values getting smaller (less correlated) as one moves from the east to the west. In SAR, the synchrony values do not appear to show as strong of an east-west spatial trend. There are also several ZIP codes with low correlation values within Central Austin. This result is not necessarily surprising given the diurnal nature of the population in these areas (i.e., large daytime, commuter populations; smaller nighttime, residential populations) and also the high population of younger individuals (at UT Austin) who may have been asymptomatic and not tested, which may have contributed to an irregular pattern of wastewater loading and case reporting.

The time lag associated with the highest correlation also varied by ZIP code, and interestingly, by sewershed. Recall that we lagged wastewater loading values against case values in 1-day intervals such that positive lag numbers indicate that SARS-CoV-2 wastewater loadings preceded COVID-19 case reporting. ZIP codes in the SAR sewershed generally have a shorter lag period as compared to those for the WC sewershed. Several ZIP codes in the SAR sewershed had a lag of 0; however, there were several other locations in SAR where wastewater loading preceded reported cases by one and four days, and one ZIP code where max synchrony occurred at a 13-day lag (78744). Also of note in the SAR sewershed were two ZIP codes with a negative lag value (78712, 78705). A negative result indicates that COVID-19 cases precede SARS-CoV-2 wastewater loadings. Although this result is possible, the likelihood that it is spurious increases due to these ZIP codes again being associated with the University of Texas at Austin which was a COVID-19 testing hub and has a large young-adult population.

In the WC service area, except for a few areas, most of the ZIP codes show a max synchrony when wastewater loading preceded positive case reports by seven days. The seven-day lag for WC departs from the aggregated analysis (Fig. 2), which found the highest correlation at a lag between zero and two days. That said, it should be noted that the synchrony values for the 1–7 day lag calculated for each ZIP code were separated by roughly a few hundredths in correlation value. It is also worthwhile to recognize the spike in correlation at the seven-day lag point for the aggregate analysis in Fig. 2. This provides some support for the 7-day finding in the disaggregated analysis.

Cluster analysis

After apportioning the daily COVID-19 time-series data, we used case rate (cases/10,000 people) to conduct a space-time cluster analysis. The analysis identified which ZIP codes were associated with COVID-19 hotspots, cold spots, and outliers as well as how long they remained part of each cluster throughout the study period. For illustrative

[5] https://www.atsdr.cdc.gov/placeandhealth/svi/index.html

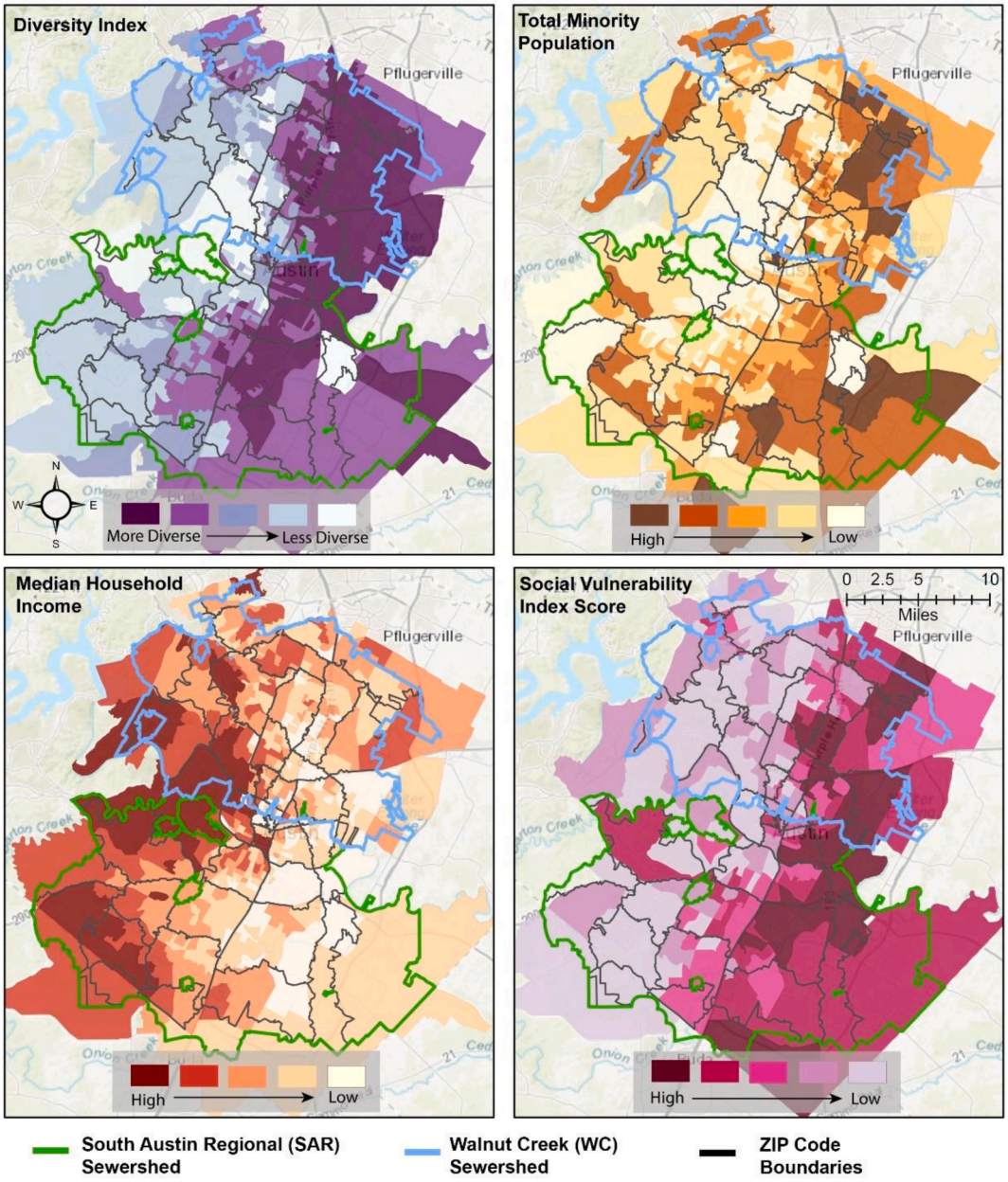


Fig. 5. Underlying socio-demographic patterns across the Austin, TX metropolitan layer with the sewershed boundaries overlain. The diversity index (top right) provides an indicator of how diverse each ZIP code is with respect to the demographic makeup of the residents. Areas with higher racial and ethnic diversity receive a higher score while areas with more homogeneity with respect to race and ethnicity receive a lower score. Total minority population (top right) measures the size of the minority population within each block group. Median household income for each block group is illustrated in the bottom left, and the social vulnerability index (top right) provides a composite measure of vulnerability based on many variables identified as corresponding to vulnerability to exogenous shocks.

Table 1
Pearson correlation values between the max synchrony values calculated for each ZIP code from the correlation between wastewater loading and new daily cases and several socioeconomic[11] indicators of interest.[22]

	Time as Hotspot	Diversity Index	Med. HH Inc.	SVI	Commuting population
Combined Max Synchrony	0.3143*	0.2085	0.1629	0.2426	.2828*
SAR Max Synchrony	0.3117	0.0636	0.3513	0.1086	.4109*
WC Max Synchrony	0.5528**	0.5124**	−0.2703	0.5219**	.1783
Significance (p-value): **<0.01, *<0.05					

purposes, we focus on hotspots. Fig. 4 displays the percentage of time that each ZIP code was part of a hotspot cluster (if at all).

In general, the pattern of COVID-19 hotspots follows an east to west trend with values high in the east and low in the west, signifying that ZIP codes in the east were a spatial and temporal hotspot for COVID-19 cases. In addition, many of the ZIP codes identified as part of a COVID-19 hotspot for a large percentage of the study period were the same ZIP codes with large synchrony values illustrated in Fig. 3.

The east-west trend in COVID-19 space-time hotspots is similar to several underlying socio-demographic patterns. To provide more nuance to these patterns, we illustrate the underlying socio-demographic patterns within Austin at the block group level in Fig. 5, with the ZIP codes outlined in dark gray. The figure highlights the stark east-west trend for most of the city. ZIP codes located in East Austin have a higher level of diversity (racial and ethnic makeup, top left panel), have a higher population of minority groups (top right), and exhibit higher social vulnerability (bottom right) as compared to West Austin. In contrast high median income areas are concentrated in West and central West Austin (bottom left panel). We also assess the spatial distribution of the percentage of the working population that commutes between 25 and 60 min for work (Appendix C) which does appear to follow a similar pattern to that of the synchrony values in Fig. 4.

In addition to the visual patterns illustrated in Fig. 5, we tested the correlation between each ZIP code's max synchrony value (the correlation between the SARS-CoV-2 wastewater loading data and new daily COVID-19 cases) and several of the socioeconomic variables and time as a hotspot (Table 1). Simultaneous analysis of ZIP codes yields two significant relationships where max synchrony is positively associated with ZIP codes that were spatial and temporal COVID-19 hotspots and the number of commuters. This result is partially expected, given that hotspots of COVID-19 are likely to be major contributors of SARS-CoV-2 levels in the wastewater and that areas with large commuting populations are positively associated with synchrony values, suggesting a relationship between commuting and viral shedding. Median household income, although not significant, shows only a very small correlation for each sewershed, although it is negative for WC (discussed in more detail below).

Interestingly, although a visual comparison of the max synchrony values for each ZIP code (Fig. 3) and the socioeconomic variables in the SAR sewershed (Fig. 5) might suggest some significant correlations, statistical testing revealed the contrary; only one of the relationships is significant. Recall that in ZIP codes with greater diversity, the index values are higher. With growing evidence to suggest that minority groups are disproportionately affected of COVID-19, we expected to see a large positive correlation between the max synchrony value and the diversity and minority index. Similarly, even though median household income was not significantly related to max synchrony in the SAR sewershed, the coefficient is positive, which is also unexpected. The coefficient for commuting population is both positive and significant, providing some evidence to suggest that commercial activity, as represented by traveling to and from work, may increase exposure and transmissivity.

The ZIP codes within the WC sewershed broadly exemplify what we expected to see across the region. Except for median household income (which is in the expected direction), all of the coefficients were significant with the expected sign. Specifically, for ZIP codes in the WC sewershed, those with a larger max synchrony value (more association between reported daily case count and wastewater loading) are positively related to the amount of time the ZIP code was a COVID-19 hotspot, a higher diversity index value, and a higher social vulnerability index value.

Discussion

The work presented here demonstrates how the disaggregation and apportioning of COVID-19 cases to smaller units of analysis can reveal more nuanced information regarding community spread when used in conjunction with the level of SARS-CoV-2 wastewater loading measured via RT-qPCR. Specifically, after apportioning cases to each respective ZIP code within the treatment plant sewershed, our analysis revealed that case counts for some ZIP codes follow the peaks and troughs of SARS-CoV-2 wastewater loading more closely than do others. When assessing in which ZIP codes the case counts best match the wastewater loading over time, we identified some spatial patterns that align with more vulnerable communities. We also found that by working at the ZIP code level rather than the aggregate sewershed area, we can forecast new case reports at an earlier time. Clinical testing for SARS-CoV-2 generally happens within 3 to 9 days after symptom onset (Freeman et al., 2021), with viral shedding beginning before symptom onset. Although the aggregated ZIP code data would provide a foreshadowing of case reports 1–2 days in advance, the ZIP code level analysis suggests that some locations wastewater loading values might precede case reports by 7–10 days. There are several implications of these findings worth further discussion.

First, we conducted cross-correlation analyses to examine the effect of disaggregating case data to a smaller unit of analysis (sewershed level vs. ZIP code level). For both analyses, we used the ratio of structures within the sewershed area for each ZIP code to total structures in order to apportion new daily cases to each ZIP code. Both analyses found that SARS-CoV-2 wastewater loading increased before the reporting of COVID-19 cases. However, there were differing temporal lags for the results. Specifically, the highest lag correlation for the aggregate analysis was between 1 and 2 days, while the highest lag correlation for the disaggregate analysis was generally between 3 and 7 days (and earlier for some), depending on ZIP code. There was a noticable second spike in correlation values for the aggregate analysis at the 7 day mark, although not as strong. These results confirm that wastewater loading is a viable approach for surveillance but that the sewershed-level (aggregate data) might not provide sufficient lead time for preparedness. Thus, the use of disaggregate case data could be better for an early-warning system.

Second, many studies have documented the positive relationship between COVID-19 case rate and vulnerable communities (DiMaggio et al., 2020; Strully et al., 2021). We know that both low-income and minority communities have a higher chance of contracting COVID-19 and suffering from COVID-19 related illness due to their higher exposure rates. Unfortunately, it is also true that testing rates are lower in these communities for a variety of reasons. This analysis offers a new perspective on understanding the geography of community spread. Using temporal cross-correlation between SARS-CoV-2 wastewater loading and COVID-19 cases, we identified the ZIP codes with the highest correlation between COVID-19 loading and daily reported cases at the ZIP code level. Our ability to assess the strength of these correlations is related to how we scaled cases. With a deeper understanding of these specific ZIP codes that are contributing more to the wastewater loading, responders can efficiently allocate human resources and response efforts. Moreover, a recent study has identified a link between the probability of contracting COVID-19 from surface water and areas where wastewater treatment plants (max capacity of >10,000 population equivalents) discharge effluent (Wang et al., 2022). The approach taken in this study could further inform this type of surface water analysis by comparing where wastewater treatment plants are discharging effluent and the case rates of the surrounding communities. This would be especially important for understanding whether vulnerable groups were disproportionately exposed to COVID-19 via wastewater treatment discharge.

Third, an important finding from this analysis was the lack of evidence (at the aggregate level) pointing to a correlation between COVID-19 cases in majority-minority communities and wastewater loading. It was only after we investigated the ZIP codes within respective sewersheds that the results matched our expectations. Still, only one of the sewersheds (WC) showed a significant correlation between max synchrony and the socioeconomic variables. These same relationships were not significant for the SAR sewershed or the region more broadly. One explanation for this result concerns the geographic concentration of minority communities in Austin. Comparing the distribution of the minority population and diversity scores at the block group level in Fig. 5 to the distribution of synchrony values in Fig. 4, two important points are revealed. First, in SAR, the concentration of minority groups within any single ZIP code is not as high compared to WC. Second, the max synchrony values associated with each ZIP code in SAR are comparatively dispersed. While there does appear to be a correlation between minority groups in WC, the case data reported at the ZIP code level does not provide us with enough granularity to determine whether the same is true in SAR. This is the likely reason why the SAR correlations in Table 1 are neither significant nor in the expected direction. That said, it is worthwhile to remember that without disaggregating the data and analyzing these relationships at the ZIP code level, we would not have been able to say anything regarding specific relationships between vulnerable communities, COVID-19 loading, and daily reported cases. As a result, the approach of case disaggregation for sewershed analyses still shows promise as a mechanism for more granular hotspot monitoring, preparedness, and mitigation decisions.

Limitations and future directions

One limitation of this work is related to the reporting of COVID-19 cases, where cases are added to the count on the COVID-19 dashboard on the date that test results become available rather than being added to the count on the date that the specimen was collected; thus, the variability with respect to the length of time for analyzing the clinical sample could affect the strength of the correlation analysis performed herein. Where the lag times are concerned, we cannot say for certain why we see some variation in the peak correlation time between wastewater loading and case reports. It may have something to do with the distance that each ZIP code is to the treatment plant but would need more data to be sure. This is a valuable avenue for future research as it may help disentangle when and to what extent specific communities are experiencing an outbreak. Another limitation of this work is the frequency of wastewater testing (two samples per week at SAR and three samples per week at WC). We also recognize that wastewater monitoring is not available nor feasible in all parts of the world, and thus there are some limitations for the generalizability of the approach detailed here. To that extent, researchers should focus on where and how to place sensors within informal sewer systems (those without a treatment plant) in order to gain insight in to where and to what extent COVID-19 is prevalent in these communities

It is also important to recognize the general pattern uncovered in this analysis compared to the location of treatment plants. With the general east-to-west trend of max synchrony values between SARS-CoV-2 wastewater loading and new daily cases and the treatment plants' location in East Austin, we cannot rule out the possibility that proximity to treatment plants had something to do with the ZIP codes that had the highest correspondence between cases and wastewater loading. Once SARS-CoV-2 enters the sewer system, it begins to degrade (Li et al., 2021); more work is required to understand the effect of signal decay and distance to a treatment plant.

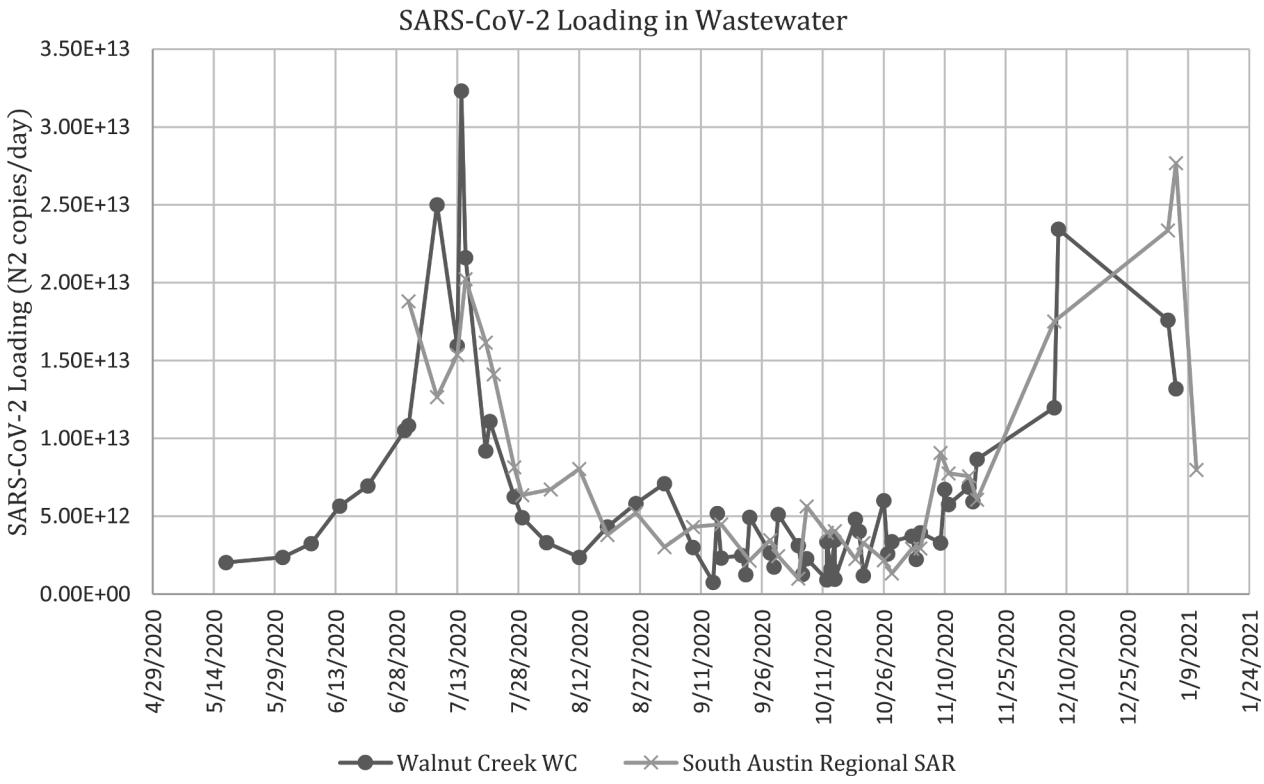
Acknowledgements

This research was funded by the Texas Division of Emergency Management (TDEM) through project number AB0718226 to MJK and KAK. Any opinions, findings, and conclusions in this article are the authors' and do not necessarily reflect the views of TDEM.We thank Austin Water personnel including Lisa Boatman, Robert Moss, Cara Cooke, Tammy West, Gary Gilmer, Trinity O'Neal, and laboratory personnel of the Walnut Creek and South Austin Regional Wastewater Treatment who provided wastewater samples, input on sampling methods, sewershed shapefiles, and treatment plant data. We thank the University of Texas at Austin personnel who helped with sample pick up (Jim Walker, Eric Kong, and Suzanne Pierce). We thank our funding sources at the University of Texas at Austin (the Fluor Centennial Teaching Fellowship in Engineering #1, the Office of the Vice President for Research at the University of Texas at Austin, and Whole Communities-Whole Health and Planet Texas 2050) and the National Science Foundation (Graduate Research Fellowship Program).

[1] HH is short for household income

[2] SVI is short for Social Vulnerability Index

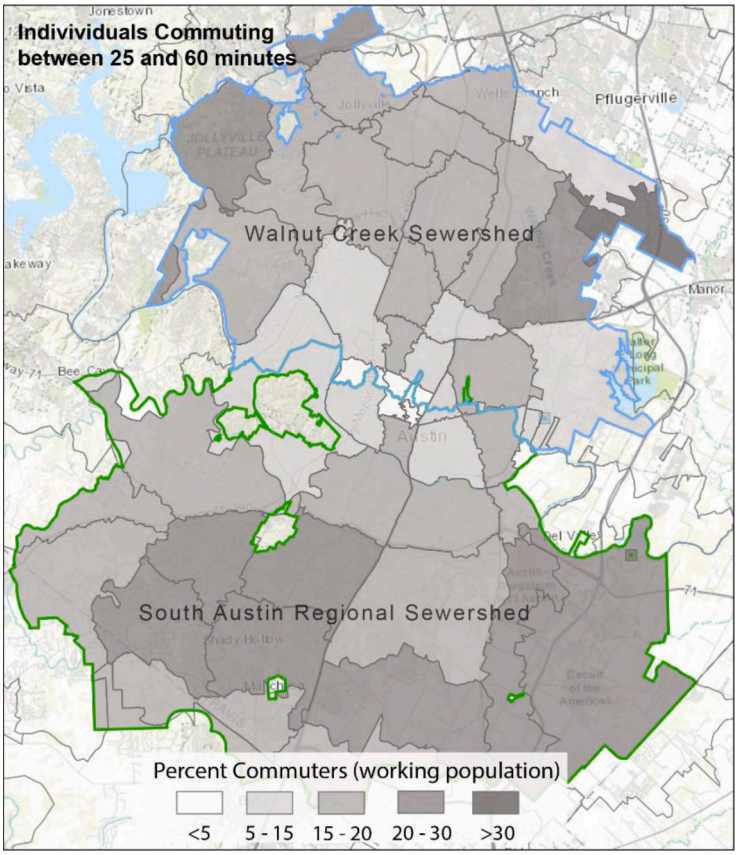
Appendix



Appendix A: Measured SARS-CoV-2 wastewater loading values overtime for each sewershed treatment plant.

B: Correlation coefficients between for each lag period when comparing the aggregate new daily reported cases and SARS-CoV2 wastewater loading.

	SAR	**WC**
Lag	**Pearson Coefficient**	**Pearson Coefficient**
0	0.6481**	0.6972**
1	0.8424**	0.6817**
2	0.4129*	0.6988**
3	0.5790**	0.5359**
4	0.5528**	0.3965**
5	0.6966**	0.3513**
6	0.6139**	0.4594**
7	0.7741**	0.5402**
8	0.7636**	0.5049**
9	0.6580**	0.4975**
10	0.5968**	0.4622**
Significance (p-value):*<0.05, **<0.01		



Appendix C: The percentage of the working population that commutes to work. Percentages are based on those workers that travel between 25 and 60 minutes to their job.

References

Ahmed, W., Angel, N., Edson, J., Bibby, K., Bivins, A., O'Brien, J.W., Choi, P.M., Kitajima, M., Simpson, S.L., Li, J., Tscharke, B., Verhagen, R., Smith, W.J.M., Zaugg, J., Dierens, L., Hugenholtz, P., Thomas, K.V., Mueller, J.F., 2020. First confirmed detection of SARS-CoV-2 in untreated wastewater in Australia: a proof of concept for the wastewater surveillance of COVID-19 in the community. Sci. Total Environ 728, 138764. https://doi.org/10.1016/j.scitotenv.2020.138764 https://doi.org/https://doi.org/.

Almagro, M., Orane-Hutchinson, A., 2020. JUE Insight: the determinants of the differential exposure to COVID-19 in New York city and their evolution over time. *J. Urban. Econ.* 103293 https://doi.org/10.1016/j.jue.2020.103293 https://doi.org/https://doi.org/.

Alwan, N.A., 2020. Surveillance is underestimating the burden of the COVID-19 pandemic. Lancet North Am. Ed 396 (10252), e24.

APH, 2021. COVID-19 Surveillance, Austin - Travis County TX. Austin Public Health. https://austin.maps.arcgis.com/apps/dashboards/39e4f8d4acb0433baae6d15a931fa984.

Balboa, S., Mauricio-Iglesias, M., Rodriguez, S., Martínez-Lamas, L., Vasallo, F.J., Regueiro, B., Lema, J.M., 2021. The fate of SARS-COV-2 in WWTPS points out the sludge line as a suitable spot for detection of COVID-19. Sci. Total Environ 772, 145268. https://doi.org/10.1016/j.scitotenv.2021.145268 https://doi.org/https://doi.org/.

Bauer, L., Broady, K., Edelberg, W., O'Donnell, J, 2020. Ten Facts about COVID-19 and the US Economy. Brookings, Washington, DC.

Bibbins-Domingo, K., 2020. This time must be different: disparities during the COVID-19 pandemic. *Ann. Intern. Med.* 173 (3), 233–234. https://doi.org/10.7326/M20-2247.

Bontempi, E., Coccia, M., Vergalli, S., Zanoletti, A., 2021. Can commercial trade represent the main indicator of the COVID-19 diffusion due to human-to-human interactions? A comparative analysis between Italy, France, and Spain. *Environ. Res.* 201, 111529 https://doi.org/10.1016/j.envres.2021.111529 https://doi.org/https://doi.org/.

British Academy. (2021). *The COVID Decade: understanding the long-term societal impacts of COVID-19.* https://www.thebritishacademy.ac.uk/documents/3238/COVID-decade-understanding-long-term-societal-impacts-COVID-19.pdf.

Caduff, C., 2020. What Went Wrong: corona and the World after the Full Stop. *Med. Anthropol. Q* 34 (4), 467–487. https://doi.org/10.1111/maq.12599 https://doi.org/https://doi.org/.

Cutter, S.L., 1996. Vulnerability to environmental hazards. *Prog. Hum. Geogr.* 20 (4), 529–539. IssueARNOLD, HODDER HEADLINE PLC.

DiMaggio, C., Klein, M., Berry, C., Frangos, S., 2020. Black/African American Communities are at highest risk of COVID-19: spatial modeling of New York City ZIP Code–level testing results. *Ann. Epidemiol.* 51, 7–13. https://doi.org/10.1016/j.annepidem.2020.08.012 https://doi.org/https://doi.org/.

Farkas, K., Hillary, L.S., Malham, S.K., McDonald, J.E., Jones, D.L., 2020. Wastewater and public health: the potential of wastewater surveillance for monitoring COVID-19. Curr. Opin. Environ. Sci. Health 17, 14–20. https://doi.org/10.1016/j.coesh.2020.06.001 https://doi.org/https://doi.org/.

Fauci, A.S., Lane, H.C., Redfield, R.R., 2020. Covid-19—Navigating the uncharted. Mass Medical Soc.

Freeman, E.E., McMahon, D.E., Hruza, G.J., Lipoff, J.B., French, L.E., Fox, L.P., Fassett, M.S., 2021. Timing of PCR and antibody testing in patients with COVID-19-associated dermatologic manifestations. *J. Am. Acad. Dermatol.* 84 (2), 505–507. https://doi.org/10.1016/j.jaad.2020.09.007.

Godri Pollitt, K.J., Peccia, J., Ko, A.I., Kaminski, N., Dela Cruz, C.S., Nebert, D.W., Reichardt, J.K.V, Thompson, D.C., Vasiliou, V, 2020. COVID-19 vulnerability: the potential impact of genetic susceptibility and airborne transmission. *Hum. Genomics* 14, 1–7.

Goldman, N., Pebley, A.R., Lee, K., Andrasfay, T., Pratt, B., 2021. Racial and ethnic differentials in COVID-19-related job exposures by occupational standing in the US. *PLoS ONE* 16 (9), e0256085. https://doi.org/10.1371/journal.pone.0256085.

Grubesic, T.H., Nelson, J.R., Wallace, D., Eason, J., Towers, S., Walker, J., 2021. Geodemographic insights on the COVID-19 pandemic in the State of Wisconsin and the role of risky facilities. *GeoJournal.* https://doi.org/10.1007/s10708-021-10503-5.

Guan, D., Wang, D., Hallegatte, S., Davis, S.J., Huo, J., Li, S., Bai, Y., Lei, T., Xue, Q., Coffman, D., 2020. Global supply-chain effects of COVID-19 control measures. Nature Hum. Behav 4 (6), 577–587.

Horowitz, J.M., Brown, A., & Minkin, R. (2021). *A year into the pandemic, long-term financial impact weighs heavily on many Americans.* https://www.pewresearch.org/social-trends/2021/03/05/a-year-into-the-pandemic-long-term-financial-impact-weighs-heavily-on-many-americans/.

Islam, M.S., 2017. Comparative evaluation of vacuum sewer and gravity sewer systems. Int. J. Syst. Assurance Eng. Manag 8 (1), 37–53. https://doi.org/10.1007/s13198-016-0518-z.

JHU, 2021. COVID-19 Dashboard by the Center for Systems Science and Engineering (CSSE) at Johns Hopkins University. Coronavirus Resource Center. https://coronavirus.jhu.edu/map.html.

Karaca-Mandic, P., Georgiou, A., Sen, S., 2021. Assessment of COVID-19 hospitalizations by Race/Ethnicity in 12 states. *JAMA Intern. Med.* 181 (1), 131–134. https://doi.org/10.1001/jamainternmed.2020.3857.

Karmakar, M., Lantz, P.M., Tipirneni, R., 2021. Association of social and demographic factors with COVID-19 incidence and death rates in the US. JAMA Netw. Open 4 (1). https://doi.org/10.1001/jamanetworkopen.2020.36462 e2036462–e2036462.

Karthikeyan, S., Ronquillo, N., Belda-Ferre, P., Alvarado, D., Javidi, T., Longhurst, C.A., Knight, R., 2021. High-Throughput wastewater sars-cov-2 detection enables forecasting of community infection dynamics in San Diego county. Msystems 6 (2) e00045-21.

Kim, H.K., Ahn, J., Atkinson, L., Kahlor, L.A., 2020. Effects of COVID-19 misinformation on information seeking, avoidance, and processing: a multicountry comparative study. *Sci Commun* 42 (5), 586–615. https://doi.org/10.1177/1075547020959670.

Kim, S.J., Bostwick, W., 2020. Social vulnerability and racial inequality in COVID-19 Deaths in chicago. Health Educ. Behav 47 (4), 509–513. https://doi.org/10.1177/1090198120929677.

Kucharski, A.J., Klepac, P., Conlan, A.J.K., Kissler, S.M., Tang, M.L., Fry, H., Gog, J.R., Edmunds, W.J., Emery, J.C., Medley, G., Munday, J.D., Russell, T.W., Leclerc, Q.J., Diamond, C., Procter, S.R., Gimma, A., Sun, F.Y., Gibbs, H.P., Rosello, A., Simons, D., 2020. Effectiveness of isolation, testing, contact tracing, and physical distancing on reducing transmission of SARS-CoV-2 in different settings: a mathematical modelling study. Lancet Infect. Dis 20 (10), 1151–1160. https://doi.org/10.1016/S1473-3099(20)30457-6.

Kumar, M., Joshi, M., Patel, A.K., Joshi, C.G., 2021. Unravelling the early warning capability of wastewater surveillance for COVID-19: a temporal study on SARS-CoV-2 RNA detection and need for the escalation. *Environ. Res.* 196, 110946.

La Rosa, G., Iaconelli, M., Mancini, P., Bonanno Ferraro, G., Veneri, C., Bonadonna, L., Lucentini, L., Suffredini, E, 2020. First detection of SARS-CoV-2 in untreated wastewaters in Italy. Sci. Total Environ 736, 139652. https://doi.org/10.1016/j.scitotenv.2020.139652 https://doi.org/https://doi.org/.

Lahrich, S., Laghrib, F., Farahi, A., Bakasse, M., Saqrane, S., El Mhammedi, M.A., 2021. Review on the contamination of wastewater by COVID-19 virus: impact and treatment. Sci. Total Environ 751, 142325. https://doi.org/10.1016/j.scitotenv.2020.142325 https://doi.org/https://doi.org/.

Larsen, D.A., Wigginton, K.R., 2020. Tracking COVID-19 with wastewater. *Nat. Biotechnol.* 38 (10), 1151–1153. https://doi.org/10.1038/s41587-020-0690-1.

Lastra, A., Botello, J., Pinilla, A., Urrutia, J.I., Canora, J., Sánchez, J., Fernández, P., Candel, F.J., Zapatero, A., Ortega, M., Flores, J., 2022. SARS-CoV-2 detection in wastewater as an early warning indicator for COVID-19 pandemic. Madrid region case study. *Environ. Res.* 203, 111852 https://doi.org/10.1016/j.envres.2021.111852 https://doi.org/https://doi.org/.

Li, H., Wang, Z., Hong, T., Piette, M.A., 2021. Energy flexibility of residential buildings: a systematic review of characterization and quantification methods and applications. Adv. Appl. Energy 3, 100054. https://doi.org/10.1016/j.adapen.2021.100054 https://doi.org/https://doi.org/.

Liao, T.F., De Maio, F., 2021. Association of social and economic inequality with coronavirus disease 2019 incidence and mortality across US counties. JAMA Netw. Open 4 (1). https://doi.org/10.1001/jamanetworkopen.2020.34578 e2034578–e2034578.

Lin Moe, T., Pathranarakul, P, 2006. An integrated approach to natural disaster management. Disaster Prevent. Manag 15 (3), 396–413. https://doi.org/10.1108/09653560610669882.

Liu, X., Feng, J., Zhang, Q., Guo, D., Zhang, L., Suo, T., Hu, W., Guo, M., Wang, X., Huang, Z., 2020. Analytical comparisons of SARS-COV-2 detection by qRT-PCR and ddPCR with multiple primer/probe sets. *Emerg Microbes Infect* 9 (1), 1175–1179.

Mahajan, U.V., Larkins-Pettigrew, M., 2020. Racial demographics and COVID-19 confirmed cases and deaths: a correlational analysis of 2886 US counties. *J. Public. Health (Bangkok)* 42 (3), 445–447. https://doi.org/10.1093/pubmed/fdaa070.

Medema, G., Heijnen, L., Elsinga, G., Italiaander, R., Brouwer, A., 2020. Presence of SARS-Coronavirus-2 RNA in sewage and correlation with reported COVID-19 prevalence in the early stage of the epidemic in The Netherlands. *Environ. Sci. Technol. Lett.* 7 (7), 511–516. https://doi.org/10.1021/acs.estlett.0c00357.

Miszta-Kruk, K., 2016. Reliability and failure rate analysis of pressure, vacuum and gravity sewer systems based on operating data. *Eng. Fail. Anal.* 61, 37–45. https://doi.org/10.1016/j.engfailanal.2015.07.034 https://doi.org/https://doi.org/.

Muñoz-Price, L.S., Nattinger, A.B., Rivera, F., Hanson, R., Gmehlin, C.G., Perez, A., Singh, S., Buchan, B.W., Ledeboer, N.A., Pezzin, L.E., 2020. Racial disparities in incidence and outcomes among patients with COVID-19. JAMA Network Open 3 (9) e2021892–e2021892.

Nemudryi, A., Nemudraia, A., Wiegand, T., Surya, K., Buyukyoruk, M., Cicha, C., Vanderwood, K.K., Wilkinson, R., Wiedenheft, B., 2020. Temporal detection and phylogenetic assessment of SARS-CoV-2 in municipal wastewater. Cell Rep. Med 1 (6), 100098.

Noh, J., Danuser, G., 2021. Estimation of the fraction of COVID-19 infected people in U. S. states and countries worldwide. *PLoS ONE* 16 (2), e0246772. https://doi.org/10.1371/journal.pone.0246772.

Palmer, E.J., Maestre, J.P., Jarma, D., Lu, A., Willmann, E., Kinney, K.A., Kirisits, M.J., 2021. Development of a reproducible method for monitoring SARS-CoV-2 in wastewater. Sci. Total Environ 799, 149405.

Petersen, E., Koopmans, M., Go, U., Hamer, D.H., Petrosillo, N., Castelli, F., Storgaard, M., Al Khalili, S., Simonsen, L., 2020. Comparing SARS-CoV-2 with SARS-CoV and influenza pandemics. Lancet Infect. Dis.

Pillay, L., Amoah, I.D., Deepnarain, N., Pillay, K., Awolusi, O.O., Kumari, S., Bux, F., 2021. Monitoring changes in COVID-19 infection using wastewater-based epidemiology: a South African perspective. Sci. Total Environ 786, 147273. https://doi.org/10.1016/j.scitotenv.2021.147273 https://doi.org/https://doi.org/.

Price-Haywood, E.G., Burton, J., Fort, D., Seoane, L., 2020. Hospitalization and mortality among black patients and white patients with Covid-19. N. Engl. J. Med 382 (26), 2534–2543. https://doi.org/10.1056/NEJMsa2011686.

Randazzo, W., Cuevas-Ferrando, E., Sanjuán, R., Domingo-Calap, P., Sánchez, G., 2020. Metropolitan wastewater analysis for COVID-19 epidemiological surveillance. *Int. J. Hyg. Environ. Health* 230, 113621. https://doi.org/10.1016/j.ijheh.2020.113621 https://doi.org/https://doi.org/.

Ranney, M.L., Griffeth, V., Jha, A.K., 2020. Critical supply shortages — the need for ventilators and personal protective equipment during the Covid-19 Pandemic. N. Engl. J. Med 382 (18), e41. https://doi.org/10.1056/NEJMp2006141.

Rodpothong, P., Auewarakul, P., 2012. Viral evolution and transmission effectiveness. *World J. Virol.* 1 (5), 131.

Saththasivam, J., El-Malah, S.S., Gomez, T.A., Jabbar, K.A., Remanan, R., Krishnankutty, A.K., Ogunbiyi, O., Rasool, K., Ashhab, S., Rashkeev, S., Bensaad, M., Ahmed, A.A., Mohamoud, Y.A., Malek, J.A., Abu Raddad, L.J., Jeremijenko, A., Abu Halaweh, H.A., Lawler, J., Mahmoud, K.A., 2021. COVID-19 (SARS-CoV-2) outbreak monitoring using wastewater-based epidemiology in Qatar. Sci. Total Environ 774, 145608. https://doi.org/10.1016/j.scitotenv.2021.145608 https://doi.org/https://doi.org/.

Schneider, D., & Harknett, K. (2021). *Essential and vulnerable: service-sector workers and paid sick leave.* https://shift.hks.harvard.edu/files/2020/04/Essential_and_Vulnerable_Service_Sector_Workers_and_Paid_Sick_Leave.pdf.

Scott, L.C., Aubee, A., Babahaji, L., Vigil, K., Tims, S., Aw, T.G., 2021. Targeted wastewater surveillance of SARS-CoV-2 on a university campus for COVID-19 outbreak detection and mitigation. *Environ. Res.* 200, 111374 https://doi.org/10.1016/j.envres.2021.111374 https://doi.org/https://doi.org/.

Strully, K., Yang, T.-C., Liu, H., 2021. Regional variation in COVID-19 disparities: connections with immigrant and Latinx communities in U.S. counties. *Ann Epidemiol* 53, 56–62. https://doi.org/10.1016/j.annepidem.2020.08.016 e2https://doi.org/https://doi.org/.

Suo, T., Liu, X., Feng, J., Guo, M., Hu, W., Guo, D., Ullah, H., Yang, Y., Zhang, Q., Wang, X., Sajid, M., Huang, Z., Deng, L., Chen, T., Liu, F., Ke, X., Liu, Y., Zhang, Q., Liu, Y., Chen, Y., 2020. ddPCR: a more sensitive and accurate tool for SARS-CoV-2 detection in low viral load specimens. MedRxiv 2020. https://doi.org/10.1101/2020.02.29.20029439, 02.29.20029439.

Tanne, J.H., 2020. Covid-19: US cases are greatly underestimated, seroprevalence studies suggest. BMJ 370. https://doi.org/10.1136/bmj.m2988 m2988.

Towers, S., Wallace, D., Walker, J., Eason, J.M., Nelson, J.R., Grubesic, T.H., 2022. A study of SARS-COV-2 outbreaks in US federal prisons: the linkage between staff, incarcerated populations, and community transmission. *BMC Public Health* 22 (1), 482. https://doi.org/10.1186/s12889-022-12813-w.

Towne, S.D., Bolin, J., Ferdinand, A., Nicklett, E.J., Smith, M.L., Ory, M.G., 2017. Assessing diabetes and factors associated with foregoing medical care among persons with diabetes: disparities facing American Indian/Alaska native, black, Hispanic, low income, and Southern Adults in the U.S. (2011–2015). *Int. J. Environ. Res. Public Health* 14 (5). https://doi.org/10.3390/ijerph14050464. Issue.

Wadhera, R.K., Wadhera, P., Gaba, P., Figueroa, J.F., Joynt Maddox, K.E., Yeh, R.W., Shen, C, 2020. Variation in COVID-19 Hospitalizations and deaths across New York City boroughs. *JAMA* 323 (21), 2192–2195. https://doi.org/10.1001/jama.2020.7197.

Wang, Z., Yang, W., Hua, P., Zhang, J., Krebs, P., 2022. Transmission risk of SARS-CoV-2 in the watershed triggered by domestic wastewater discharge. Sci. Total Environ 806, 150888. https://doi.org/10.1016/j.scitotenv.2021.150888 https://doi.org/https://doi.org/.

Weichselgartner, J., 2001. Disaster mitigation: the concept of vulnerability revisited. Disaster Prevent. Manag 10 (2), 85–95.

Weidhaas, J., Aanderud, Z.T., Roper, D.K., VanDerslice, J., Gaddis, E.B., Ostermiller, J., Hoffman, K., Jamal, R., Heck, P., Zhang, Y., 2021. Correlation of SARS-CoV-2 RNA in wastewater with COVID-19 disease burden in sewersheds. Sci. Total Environ 775, 145790.

WHO. (2020). *Coronavirus disease 2019 (COVID-19): situation report 82.* https://apps.who.int/iris/bitstream/handle/10665/331780/nCoVsitrep11Apr2020-eng.pdf.

Wu, S.L., Mertens, A.N., Crider, Y.S., Nguyen, A., Pokpongkiat, N.N., Djajadi, S., Seth, A., Hsiang, M.S., Colford, J.M., Reingold, A., Arnold, B.F., Hubbard, A., Benjamin-Chung, J., 2020. Substantial underestimation of SARS-CoV-2 infection in the United States. *Nat. Commun.* 11 (1), 4507. https://doi.org/10.1038/s41467-020-18272-4.

Yeager, R., Holm, R.H., Saurabh, K., Fuqua, J.L., Talley, D., Bhatnagar, A., Smith, T., 2021. Wastewater sample site selection to estimate geographically-resolved community prevalence of COVID-19: a sampling protocol perspective. GeoHealth 5 e2021GH000420.

Zarefsky, M., 2020. Why African American communities are being hit hard by COVID-19. Am. Med. Assoc 1 https://www.ama-assn.org/delivering-care/population-care/why-african-american-communities-are-being-hit-hard-covid-19.