International Journal of Applied Earth Observations and Geoinformation 112 (2022) 102855

Contents lists available at ScienceDirect

International Journal of Applied Earth Observations and Geoinformation

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

Associations between nighttime light and COVID-19 incidence and mortality in the United States

Yiming Zhang [a,1], Ningyezi Peng [a,b,1], Shujuan Yang [c,a,*], Peng Jia [a,d,*]

[a] International Institute of Spatial Lifecourse Health (ISLE), Wuhan University, Wuhan, China
[b] Department of Land Surveying and Geo-Informatics, The Hong Kong Polytechnic University, Hong Kong, China
[c] West China School of Public Health and West China Fourth Hospital, Sichuan University, Chengdu, China
[d] School of Resource and Environmental Sciences, Wuhan University, Wuhan, China

ARTICLE INFO

Keywords:
COVID-19
Pandemic
Epidemic
Nighttime light
Spatial epidemiology

ABSTRACT

COVID-19 has caused almost 770,000 deaths in the United States by November 2021. The nighttime light (NTL), representing the intensity of human activities, may reflect the degree of human contacts and therefore the intensity of COVID-19 transmission. This study intended to assess the associations between NTL differences and COVID-19 incidence and mortality among U.S. counties. The COVID-19 data of U.S. counties as of 31 December 2020 were collected. The average NTL values for each county in 2019 and 2020 were derived from satellite data. A negative binomial mixed model was adopted to assess the relationships between NTL intensity and COVID-19 incidence and mortality. Compared to the counties with the lowest NTL level (0.14–0.37 $nW/cm^2/sr$), those with the highest NTL level (1.78–59.61 $nW/cm^2/sr$) were related with 15% higher mortality rates (mortality rate ratio:1.15, 95 %CI: 1.02–1.30, p-value: 0.02) and 23% higher incidence rates (incidence rate ratio:1.23, 95 %CI: 1.13–1.34, p-value < 0.0001). Our study suggested that more intensive NTL was related with higher incidence and mortality rates of COVID-19, and NTL had a stronger correlation with the COVID-19 incidence rate than mortality rate. Our findings have contributed solid epidemiological evidence to the existing COVID-19 knowledge pool, and would help policymakers develop interventions when faced with the potential risk of the following outbreaks.

1. Introduction

COVID-19 has spread across countries and been seen as a global emergency since January 30, 2020 (Jia, 2020). After more than one year, the transmission of COVID-19 remains a grave threat to global public health, especially in the U.S. By the end of June 2021, there have been the largest number of COVID-19 infection and death cases in the U. S. (Johns Hopkins Coronavirus Resource Center, 2021), with nearly every county affected (Centers for Disease Control and Prevention, 2021). Due to the limited efficacy of COVID-19 vaccines to virus variants (Lopez Bernal et al., 2021), social distancing is still the most important and effective measure to reduce disease transmission by restricting human interactions. Some environmental factors, such as greenness (Klompmaker et al., 2021), air quality (Wu et al., 2020), temperature, and humidity (Ma, 2020), have been found to be indirectly related to the COVID-19 transmission process. For example, from a data-driven perspective, green space may serve as relatively safe places for social interactions while not increasing the risk of infection in densely populated counties in the U.S. (Klompmaker et al., 2021). However, such environmental factors cannot directly reflect the intensity of human contacts which is necessary to the outbreak and transmission of infectious diseases.

Different from those environmental factors, nighttime light (NTL) can serve as direct evidence of the intensity of human contacts (Huang et al., 2014). The NTL satellite imagery has been utilized for estimating population density (Sutton and Obremski, 2003; Sutton et al., 2001), measuring and mapping Gross Domestic Product (GDP) at national and sub-national levels (Ebener, 2005; Henderson et al., 2012; Mayuri Chaturvedi et al., 2011; Sutton, 2007). According to previous studies, NTL intensity could reflect the extents of urbanization (Zhang and Seto,

* Corresponding authors at: West China School of Public Health and West China Fourth Hospital, Sichuan University, Chengdu, China (S. Yang); School of Resource and Environmental Sciences, Wuhan University, Wuhan, China (P. Jia).

E-mail addresses: rekiny@126.com (S. Yang), jiapengff@hotmail.com, jiapeng@whu.edu.cn (P. Jia).

[1] Equal contribution.

https://doi.org/10.1016/j.jag.2022.102855
Received 2 November 2021; Received in revised form 28 May 2022; Accepted 30 May 2022
Available online 18 June 2022

2011), economic growth (Henderson et al., 2012), global poverty (Elvidge et al., 2009), etc., which may have direct or indirect impacts on COVID-19 transmission. Moreover, individuals' exposure to more intensive NTL has been associated with higher odds of obesity (Zhang et al., 2020), and people with obesity may be more susceptible to severe COVID-19 outcomes (CDC, 2021). Some studies have examined the changes of NTL during the periods of social distancing in several megacities, such as Wuhan in China and Los Angeles and New York in the U.S. (Liu et al., 2020; Xu, 2021). However, to our knowledge, few studies have statistically linked the changes of NTL with both COVID-19 incidence and mortality together.

This study intended to investigate the associations between NTL and COVID-19 incidence and mortality in the contiguous U.S. We first calculated the mean NTL values within each U.S. county in 2019 and 2020 separately and compared their changes. Then, we adopted a negative binomial mixed model to explore the relationships between NTL and COVID-19 outcomes, adjusting for comprehensive covariates. Furthermore, we conducted sensitivity analysis by adding order (stay at home/shelter to end/relax stay at home/shelter in place) issuance days and the state-level number of COVID-19 tests. The results can reflect the impact of human activities on COVID-19 transmission, which could support decision making in controlling the following outbreaks of the pandemic and post-pandemic reopening strategies in the U.S. and even in other countries of the world.

2. Methods

2.1. Exposure variable

The monthly cloud-free NTL images for the contiguous U.S. in 2019 and 2020 at 15-arc second resolution (about 450 m) were derived from nighttime Visible Infrared Imaging Radiometer Suite Day Night Band (VIIR/DNB) on the Joint Polar-orbiting Satellite System, provided by the Earth Observation Group of Payne Institute for Public Policy. Compared with the commonly used Operational Linescan Sensor (OLS) onboard Defense Meteorological Satellite Program (DMSP) satellites, the VIIRS/DNB has significant improvement on the characterization of low levels of nighttime light (Earth Observation Group; Elvidge et al., 2013; Elvidge et al., 2017; Elvidge et al., 2021), by excluding background noise, solar and lunar contamination, data degraded by cloud cover, and features unrelated to electric lighting (Elvidge et al., 2017). The annual mean NTL value in each year was calculated for each U.S. county, to represent the average intensity of NTL over the year and compare the NTL changes between the two years.

2.2. Outcome variables

We collected the daily numbers of infected and death cases of COVID-19 in each U.S. county in 2020, from the U.S. Center for Disease Control and Prevention (CDC) and the Center for Systems Science and Engineering (CSSE) of Johns Hopkins University (Dong et al., 2020), and aggregated them at the county level.

2.3. Covariates

Environmental factors, including particulate matter with aerodynamic diameter < 2.5 μm ($PM_{2.5}$), temperature, and humidity (Wu et al., 2020; Ma, 2020), were adjusted in the study. We calculated the summertime and wintertime averages of temperature and relative humidity data from 2000 to 2016 for each U.S. county based on GridMET data (Abatzoglou, 2013). We made county-level $PM_{2.5}$ exposure data from 2000 to 2016 by an established $PM_{2.5}$ prediction model (Aaron, et al., 2019).

Eleven county-level census variables, which might be associated with NTL and COVID-19 outcome (Klompmaker et al., 2021; Wu et al., 2020), were also obtained from the 2000–2018 decennial census, the America Community Survey (ACS) (The United States Census Bureau and Data, 2021) and the 2009–2016 CDC Mortality File (Centers for Disease Control and Prevention), and included in our analysis to adjust for potential confounders: proportion of inhabitants in poverty, population density, median household value, percentage of Black inhabitants, percentage of Hispanic inhabitants, percentage of residents with a secondary school certificate, median household income, percentage of inhabitants who own their house, percentage of inhabitants older than 65, percentage of inhabitants aged 45–64, percentage of inhabitants aged 15–44.

The COVID-19 outbreaks and response and the accessibility of health services in counties may also confound the connection between NTL and COVID-19 outcome. We also calculated the days from the issuance of stay at home/shelter to end/relax these policies to assess whether the association between NTL and COVID-19 would differ by the state government policy (Raifman, et al., 2020). We counted the number of days from the first recorded COVID-19 case to represent the extent of the COVID-19 pandemic. We collected the number of hospital beds (county-level) in 2019 and COVID-19 tests (state-level, up to December 31, 2020) to evaluate the influence of public sanitation resources on the analysis (The COVID Tracking Project. 2021, HIFLD). We additionally included two county-level behavioral risk variables, namely, the percentage of inhabitants with obesity and the percentage of current smokers, measured in 2011 and obtained from the U.S. CDC Behavioral Risk Factor Surveillance System (BRFSS).

2.4. Statistical analysis

A negative binomial mixed model (Zhang et al., 2017; Booth et al., 2003), which can effectively manage the over-dispersion of the longitudinal data (Yau et al., 2003), was adopted to assess the association between NTL and the incidence and mortality of COVID-19. All U.S. counties were divided into quintiles by the annual mean NTL value, with the 1st quintile (the lowest NTL level) as the reference group in the model. We took NTL as the exposure of interest and reported COVID-19 incidence rate ratios (IRR) and mortality rate ratios (MRR) as the outcome to estimate the association, adjusted by covariates. To assess the effects of potential confounders, we conducted a series of models by adjusting covariates step by step. Specifically, Model 1 only included a population size offset and a random intercept by state, which was carried out to interpret for potential correlation in counties due to similar COVID-19 response policies in the same state. All census and BRFSS covariates were additionally adjusted in Model 2. In Model 3, we further adjusted the days from the first recorded COVID-19 case and the number of hospital beds. In Model 4, we further adjusted temperature, humidity, and $PM_{2.5}$. The IRR and MRR for the n-th quintile was considered the increase in the COVID-19 incidence and mortality rates related with the relative variety from the first quintile to the n-th quintile in NTL exposure. We hypothesized that association with NTL and COVID-19 incidence and mortality was stronger in high levels of NTL exposure. All analyses were carried out in R software and performed model fitting using the lme4 package (Bates et al., 2015). The main model takes the following form (E[·] denotes an expected value, taking IRR as an example):

$$\log[E(\text{COVID-19 cases})] = \beta_0 + \beta_1 \text{ NTL} + \beta_2 \text{ \% living in poverty} + \beta_3 \text{ population density} + \beta_4 \text{ median household value} + \beta_5 \text{ \% Black} + \beta_6 \text{ \% Hispanic} + \beta_7 \text{ \% adults with less than a high school education} + \beta_8 \text{ median household income} + \beta_9 \text{ \% owner-occupied housing} + \beta_{10} \text{ \% inhabitants} \geq 65 + \beta_{11} \text{ \% inhabitants aged 45–64} + \beta_{12} \text{ \% inhabitants aged 15–44} + \beta_{13} \text{ \% obesity} + \beta_{14} \text{ \% current smokers} + \beta_{15} \text{ days since first case} + \beta_{16} \text{ number of hospital beds} + \beta_{17} PM_{2.5} + \beta_{18} \text{ average summer temperature} + \beta_{19} \text{ average winter temperature} + \beta_{20} \text{ average winter relative humidity} + \beta_{21} \text{ average summer relative humidity} + \text{offset}[\log(\text{population size})] + \text{random intercept (state)}.$$

We reported the IRR and MRR and 95% confidence intervals (CIs) for NTL, corresponding to the exponentiated parameter estimate ($e^{\beta 1}$).

To assess the robustness of the result, several sensitivity analyses were conducted. First, we added the days from the release of the stay at home/shelter in place to the end of the order to evaluate the impact of state response in the main model. Second, we added the state-level number of COVID-19 tests to the main model to assess the impact of public sanitation. Third, we added both variables to the main model.

3. Results

3.1. Basic characteristics of the U.S. counties

The basic characteristics of the 3,134 U.S. counties were presented by quintile (Q1-Q5) in Table 1. Among them, 3,106 counties reported 10 or more COVID-19 cases. Compared to 2019, the mean county-level NTL value in highest quintiles (Q5) in 2020 was significantly decreased, reaching 6% in Q5 (6.31 in 2019 and 5.94 in 2020). Counties in Q3 had the highest incidence rates and mortality rates occurred (IR = 68,280.03, MR = 1,117.98), compared to those in other quintiles. In addition, counties with higher NTL had more Blacks and the higher population density. The counties with higher NTL had more residents aged 15–44 and less aged 45–64 and over 65. The counties in Q3 and Q4 had a higher proportion of poverty, residents with obesity, and current smokers.

There were spatial heterogeneities in the NTL level (Fig. 1) and COVID-19 incidence and mortality rates (Fig. 2) across the U.S. The counties with higher NTL values in 2020 were mainly in the midwest, northeast, and south of the U.S. (Fig. 1**a**), and the states on the east coast had more counties with NTL in high quintiles (Q4: 0.84–1.65 and Q5: 1.66–59.61). However, from east to west, the number of counties with high-level NTL was decreasing. Moreover, in the central region, there were only several states that had counties with NTL in high quintiles (e. g., North Dakota, Colorado, New Mexico). To the west, counties with high-level NTL were mainly in California and Washington. Also, changes in NTL from 2019 to 2020 varied across the U.S. (Fig. 1**b**). Due to lockdown and policies that responded to the COVID-19, the mean NTL of most counties decreased in 2020, especially in the western and southern counties, but NTL values of midwestern and northeastern counties had slightly increased in 2020. As for COVID-19 distributions, the counties with the highest incidence rates (Q5) were mainly in midwestern and southern regions of the U.S., and counties in northeastern and central U. S. had the highest COVID-19 mortality rates (Q5) to the end of 2020 (Fig. 2).

3.2. Associations between NTL and COVID-19 incidence and mortality

Compared to the counties in Q1 (reference group), the counties in Q5 showed 23% higher incidence rate (IRR=1.23 [95% CI: 1.13–1.34]) and 15% higher mortality rate (MRR=1.15 [95% CI: 1.02–1.30]) (Table 2). The counties in Q4 showed 19% higher incidence rate (IRR=1.19 [95% CI: 1.11–1.27]) and 12% higher mortality rate (MRR=1.12 [95% CI: 1.02–1.24]). The counties in Q2 and Q3 showed only 9% (IRR=1.09 [95% CI: 1.03–1.14]) and 15% higher incidence rates (IRR=1.15 [95% CI: 1.09–1.23]), respectively, without showing the significantly higher mortality rates ($MRR_{Q2\ vs\ Q1}$ = 1.01 [95% CI: 0.93–1.09] and $MRR_{Q3\ vs\ Q1}$ = 1.09 [95% CI: 0.99–1.19]).

In sensitivity analyses, by adding order issuance days, state-level COVID-19 tests, and both variables to the models, we found consistent associations between NTL and COVID-19 incidence and mortality rates (Table 3).

4. Discussion

Many existing epidemiological researches have implied that environmental factors, such as temperature, air pollution, humidity, and outdoor greenness, may influence the spread of COVID-19 because virus stability and host susceptibility, and contact rates can be affected by them (Klompmaker et al., 2021; Notari, 2021; Wang, 2021; Wu et al., 2020). However, to our knowledge, few studies have statistically linked the changes of NTL with both the incidence and mortality of COVID-19 together. In this study, we analyzed the associations between NTL and COVID-19 outcomes, covering 3134 counties in the U.S and the full year time in 2020. We observed that NTL was positively related with both COVID-19 incidence and mortality, and NTL had stronger associations with incidence rates than mortality rates.

Table 1
Descriptive statistics of the 3,134 U.S. counties by quintiles of nighttime light (NTL).

Variables	Q1	Q2	Q3	Q4	Q5
NTL					
NTL in 2020, nW/cm^2/sr, mean (range)	0.28 (0.14–0.36)	0.46 (0.36–0.56)	0.69 (0.56–0.84)	1.14 (0.84–1.65)	5.94 (1.66–59.61)
NTL in 2019, nW/cm^2/sr, mean (range)	0.26 (0.14–0.44)	0.44 (0.28–1.20)	0.68 (0.32–1.59)	1.15 (0.68–1.95)	6.31 (1.55–68.60)
COVID-19 outcome					
Incidence rate (per million) (median)	59,765.21	65,729.87	68,280.03	66,935.85	63,352.25
Mortality rate (per million) (median)	791.45	999.17	1117.98	1076.60	1016.90
Meteorological covariates					
Winter temperature, K	276.49	279.30	281.40	282.64	281.90
Summer temperature, K	301.63	303.11	303.89	304.06	303.82
Winter relative humidity, %	89.23	88.92	88.03	87.62	86.27
Summer relative humidity, %	86.12	91.37	92.86	92.97	91.36
County-level covariates					
Poverty, %	8.92	9.35	9.96	9.53	8.12
Population density, people/sq.mi.	5.77	25.70	53.00	114.88	765.40
Median house value, $	109,062.50	94,114.29	98,550.00	110266.70	155,220.50
Black inhabitants, %	0.28	0.66	1.58	2.31	7.46
Hispanic inhabitants, %	2.99	2.35	2.350	3.09	5.70
Education, %	14.63	19.16	22.57	22.06	18.25
Household income, $	46,275.00	46,051.83	45,525.50	46,908.70	56,333.98
Owner occupied house, %	77.26	78.90	77.47	76.55	71.11
Aged over 65, %	19.44	17.19	15.74	14.57	12.65
Aged 45–64, %	27.99	27.04	26.49	26.30	25.29
Aged 15–44, %	33.27	36.18	37.77	38.88	41.15
Obesity, %	30.80	33.20	34.60	34.40	31.50
Current smoker, %	15.34	16.97	18.32	18.23	16.62
$PM_{2.5}$, μg/m^3	4.79	7.96	9.30	10.00	10.67
Rate of hospitals beds (per million)	1,968.43	1,693.77	1,767.46	2,217.82	3,066.40

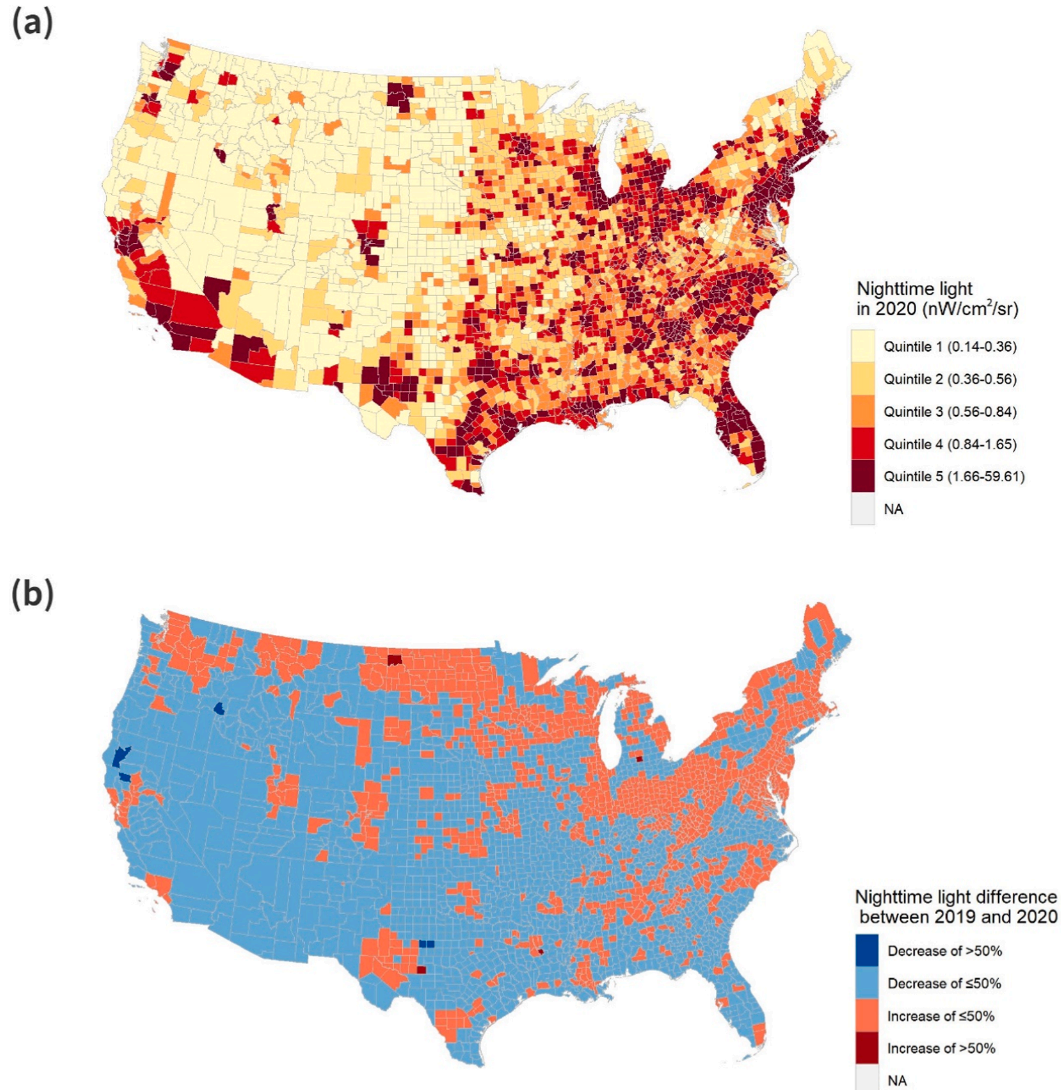

Fig. 1. Spatial distribution of nighttime light in 2020 (a) and the differences in nighttime light between 2019 and 2020 (b) within each U.S. county.

The county-level NTL values in the U.S. were decreased during the COVID-19 period than previously, which was consistent with the NTL change in China (Liu et al., 2020). Population density, economic activity, and state-level policy response to COVID-19 may affect the mean NTL value of counties in 2020 (Fig. 1**a**) and NTL difference between 2019 and 2020 (Fig. 1**b**). Studies had reported that NTL was corresponded with population density (Sutton et al., 2001). The same as shown in Table 1, the population density increased by multiples with the change of NTL quintiles. For example, compared to Q1 (5.77 people/square miles), the population density in counties within Q5 has increased nearly 132 times to 765.40 people/square miles. In addition, NTL had a close relationship with economic activity (Henderson et al., 2012; Mayuri Chaturvedi et al., 2011). The real GDP in the third and fourth quarter in 2020 increased 33.4% and 4.3%, respectively, and the real GDP of almost all U.S. states increased in the 2020 final quarter (The U.S. Bureau of Economic Analysis, 2020). This may account for the slight increase in NTL in some counties compared to 2019, even though the national annual real GDP reduced 3.5% in 2020 compared to 2019. Another reason for the increased NTL in some counties could be the difference in the state-level policy response to COVID-19. North Dakota, for example, did not execute the stay-at-home/shelter in place order (Raifman, et al., 2020).

We found that compare to the first quintile of NTL value, the fifth quintile was associated with 15% and 23% higher incidence and mortality rates, respectively, of COVID-19. Our results of incidents rates were consistent with Meng, Zhu (Meng et al., 2021), which reported a notable positive relationship between NTL intensity and COVID-19 incidence rates in New York and Connecticut. Different from Meng, Zhu (Meng et al., 2021), we also studied the relationship between NTL and COVID-19 mortality rates, and our results included all counties in the U.S. and covered a more extended period. According to a priori knowledge, NTL can be used as a proxy for investigating diverse anthropogenic activities from the regional to the country scale (Ma et al., 2018), and zones of high economic activity were highlighted in NTL imagery (Kyba et al., 2015). In response to the COVID-19 outbreak, policies, such as stay at home and non-essential business closure, had restricted many kinds of anthropogenic activities, and thus influenced NTL distribution. Therefore, NTL can also serve as a proxy for human contacts which may further impact COVID-19 transmission (Small and Sousa, 2021). Counties with NTL in high quintiles may be related to higher economic activities, while higher economic activities increased the risk of exposure to COVID-19. For example, air travel contributed to the transmission of COVID-19 throughout the world and accelerated the global pandemic (Murphy et al., 2020); the transportation of cold-chain food has been confirmed as a way of transmission COVID-19 (Ji et al., 2021); exposed public transportation also increased attack rate of

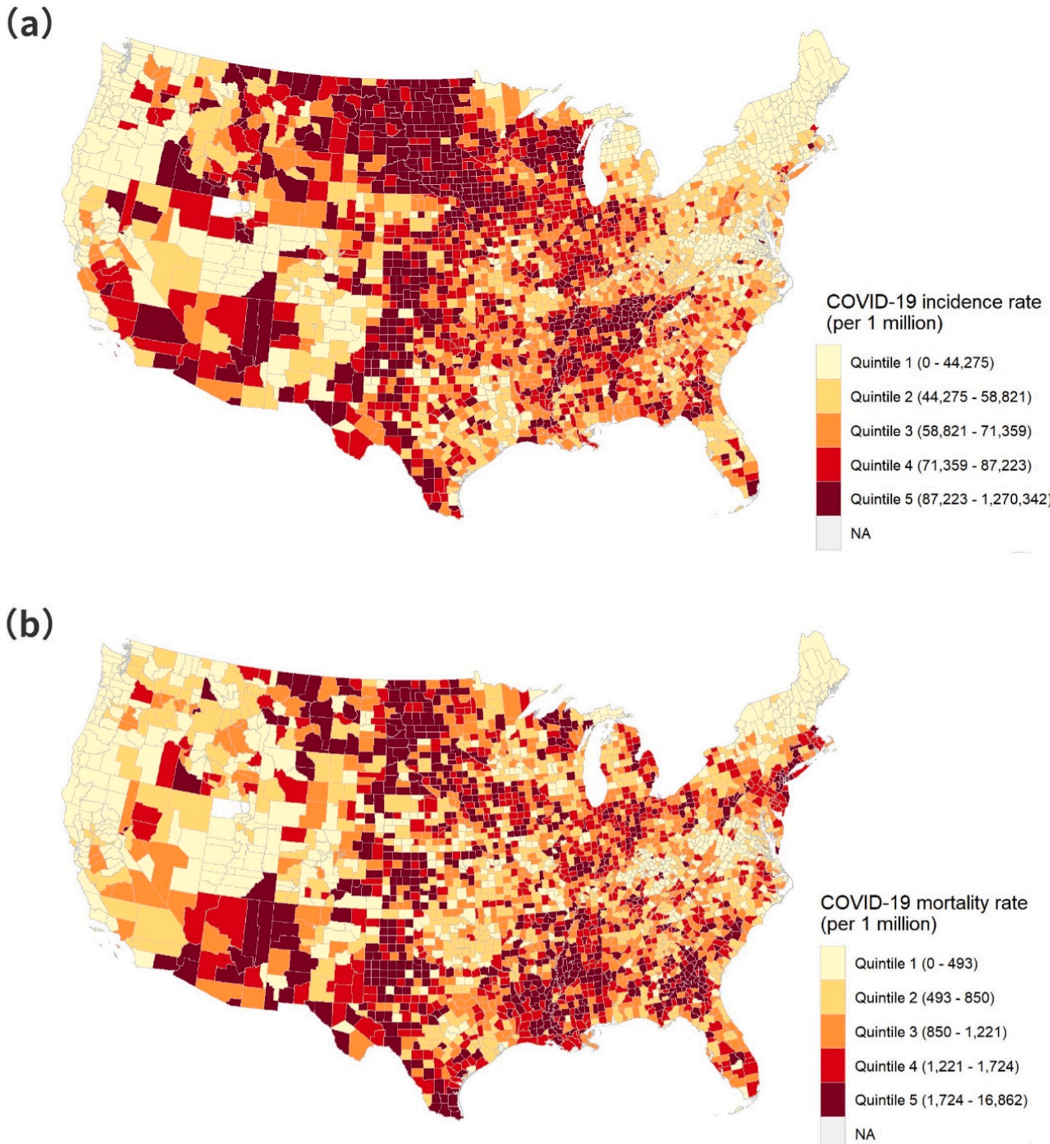

Fig. 2. County-level COVID-19 incidence rate (a) and mortality rate (b) in the U.S.

Table 2
Associations of nighttime light (NTL) with COVID-19 incidence and mortality rates.

	COVID-19 incidence rate			COVID-19 mortality rate		
NTL in quintile	Incidence rate ratio	95% CI	p-value	Mortality rate ratio	95% CI	p-value
Q1	Ref			Ref		
Q2	1.09	1.03–1.14	0.002	1.01	0.93–1.09	0.88
Q3	1.15	1.09–1.23	<0.0001	1.09	0.99–1.19	0.07
Q4	1.19	1.11–1.27	<0.0001	1.12	1.02–1.24	0.02
Q5	1.23	1.13–1.34	<0.0001	1.15	1.02–1.30	0.02

COVID-19 (Shen, 2020). This may account for why the higher quintile NTL was associated with higher rates in incidence and mortality in COVID-19 compared to the lowest quintile of NTL.

The results indicated that NTL had stronger associations with COVID-19 incidence rates than mortality rates, perhaps due to the time span between infection and possible death. Furthermore, some deaths due to COVID-19 complications may not be counted as COVID-19 deaths. A study suggested that COVID-19 patients who got infected have a 59% higher risk of death up to 6 months after infection than non-infected people, and there were eight additional deaths every 1,000 patients within six months due to long-term COVID-19 complications were not counted as COVID-19 deaths (Al-Aly et al., 2021). In lower quintile (Q2 and Q3), NTL has little association with mortality rates, and neither of the p-values was statistically significant.. Counting errors in the number of COVID-19 deaths may be responsible for the little associations between COVID-19 mortality rates and NTL in the lower quintile of NTL. Deaths through long-term COVID-19 complications were not counted as COVID-19 deaths (Al-Aly et al., 2021). And counties with lower quintiles NTL had more residents aged over 65 (Table 1), and elderly people with conditions could be more inclined to have severe

Table 3
Associations of nighttime light and COVID-19 incidence and mortality rates in sensitivity analyses.

	NTL in quintile	COVID-19 incidence rate Incidence rate ratio	95% CI	p-value	COVID-19 mortality rate Mortality rate ratio	95% CI	p-value
Adding order issuance days	Q1	Ref			Ref		
	Q2	1.09	1.03–1.14	0.002	1.01	0.93–1.09	0.88
	Q3	1.15	1.09–1.23	<0.001	1.09	0.99–1.19	0.07
	Q4	1.19	1.11–1.27	<0.001	1.12	1.02–1.24	0.02
	Q5	1.23	1.13–1.34	<0.001	1.15	1.02–1.30	0.02
Adding state-level number of COVID-19 tests	Q1	Ref			Ref		
	Q2	1.09	1.03–1.14	0.002	1.01	0.93–1.09	0.88
	Q3	1.15	1.09–1.23	<0.001	1.09	0.99–1.19	0.07
	Q4	1.19	1.11–1.27	<0.001	1.13	1.02–1.24	0.02
	Q5	1.23	1.13–1.34	<0.001	1.15	1.02–1.30	0.02
Adding both variables above	Q1	Ref			Ref		
	Q2	1.09	1.03–1.14	0.002	1.01	0.93–1.09	0.87
	Q3	1.15	1.09–1.23	<0.001	1.09	0.99–1.19	0.07
	Q4	1.19	1.11–1.27	<0.001	1.12	1.02–1.24	0.02
	Q5	1.23	1.13–1.34	<0.001	1.15	1.02–1.30	0.02

COVID-19 symptoms (CDC, 2021).

To evaluate the robustness of our results, we conducted several sensitivity analyses by adjusting for more COVID-related variables. A strong relationship has been found between the issuance date of stay-at-home orders and COVID-19 incidence during the early stage of COVID-19 (Klompmaker et al., 2021). We used the effective days of the stay-at-home order by the end of 2020 as the proxy of the state response to COVID-19 (January 25, 2021, for California). After adjusting for the stay-at-home orders in the sensitivity analysis, the fifth quintile was associated with 15% and 23% higher incidence and mortality in COVID-19 than the lowest quintile, which was consistent with the main model.

This study has some major limitations. First, an ecological study design at the county level, although cost-effective and straightforward to analyze the association between NTL and COVID-19 incidence and mortality, may not reveal individual risk factors due to the ecological fallacy, which leaves us incapable to make inferences of individual-level associations (Jia, 2019). Second, NTL derived from satellite imagery may merely be a proxy measure of the NTL exposure from outdoor rather than indoor nightlight (Jia et al., 2019), which may bring about misclassification in NTL exposure (Zhang et al., 2020; Jia et al., 2020). The number of underreported COVID-19 cases may also influence our analysis. Third, like most of epidemiological studies, this study aimed to examine the associations between NTL and COVID-19 incidence and mortality rates, rather than predicting COVID-19 incidence and mortality rates. Therefore, the associations found in this study are to be validated in future research which aims at COVID-related prediction.

5. Conclusions

This study suggests that more intensive NTL may be associated with the higher COVID-19 incidence and mortality rates, and NTL may be more strongly associated with COVID-19 incidence rate than mortality rate. A total of 3,134 U.S. counties and 98% of the U.S. population are covered in this study, and the NTL data for the entire year of 2020 were used, overlapping with different stages of the COVID-19 pandemic in the U.S. Adjusting for a wide range of covariates obtained from multiple sources and additional sensitivity analysis further consolidate the robustness of our models. Our findings have added solid epidemiological evidence to the existing COVID-19 knowledge pool, and would help policy-makers develop interventions when faced with the potential risk of the following outbreaks.

Disclosure statement.

All authors have nothing to disclose.

CRediT authorship contribution statement

Yiming Zhang: Data curation, Methodology, Formal analysis, Visualization, Writing – original draft. **Ningyezi Peng:** Data curation, Methodology, Formal analysis, Visualization, Writing – original draft. **Shujuan Yang:** Methodology, Writing – original draft, Supervision. **Peng Jia:** Conceptualization, Methodology, Writing – original draft, Supervision.

Declaration of Competing Interest

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

Acknowledgement

We thank Wuhan University Specific Fund for Major School-level Internationalization Initiatives (WHU-GJZDZX-PT07) and the International Institute of Spatial Lifecourse Health (ISLE) for the research support.

References

Aaron, v.D., et al., 2019. *Regional Estimates of Chemical Composition of Fine Particulate Matter Using a Combined Geoscience-Statistical Method with Information from Satellites, Models, and Monitors.* Environmental Science & Technology, **53**(5): p. 2595-2611.

Abatzoglou, J.T., 2013. Development of gridded surface meteorological data for ecological applications and modelling. Int. J. Climatol. 33 (1), 121–131.

Al-Aly, Z., Xie, Y., Bowe, B., 2021. High-dimensional characterization of post-acute sequelae of COVID-19. Nature 594 (7862), 259–264.

Bates, D., Machler, M., Bolker, B., Walker, S., 2015. *Fitting linear mixed-effects models using lme4.* J. Stat.Softw. **67**: p. 1–48.

Booth, J.G., Casella, G., Friedl, H., Hobert, J.P., 2003. *Negative binomial loglinear mixed models.* Stat. Model. **3**: p. 179–191.

Centers for Disease Control and Prevention, *Data Access - Compressed Mortality File.*

Centers for Disease Control and Prevention. *CDC COVID Data Tracker*. 2021; Available from: https://covid.cdc.gov/covid-data-tracker/#county-view.

CDC. *Certain Medical Conditions and Risk for Severe COVID-19 Illness.* [cited 2021 June 02]; Available from: https://www.cdc.gov/coronavirus/2019-ncov/need-extra-precautions/people-with-medical-conditions.html.

Dong, E., Du, H., Gardner, L., 2020. An interactive web-based dashboard to track COVID-19 in real time. Lancet Infect. Dis. 20 (5), 533–534.

Earth Observation Group. *See the World at Night.* [cited 2021 June 02]; Available from: https://eogdata.mines.edu/products/vnl/.

Ebener, S., et al., 2005. From wealth to health: modelling the distribution of income per capita at the sub-national level using night-time light imagery. Int. J. Health Geographics 4 (1), 5.

Elvidge, C.D., Sutton, P.C., Ghosh, T., Tuttle, B.T., Baugh, K.E., Bhaduri, B., Bright, E., 2009. A global poverty map derived from satellite data. Comput. Geosci. 35 (8), 1652–1660.

Elvidge, C.D., Baugh, K.E., Zhizhin, M., Hsu, F.-C., 2013. Why VIIRS data are superior to DMSP for mapping nighttime lights. *Why VIIRS data are superior to DMSP for mapping nighttime lights* Asia-Pacific Advanced Network 35 (0), 62.

Elvidge, C.D., Baugh, K., Zhizhin, M., Hsu, F.C., Ghosh, T., 2017. VIIRS night-time lights. Int. J. Remote Sens. 38 (21), 5860–5879.

Elvidge, C.D., Zhizhin, M., Ghosh, T., Hsu, F.-C., Taneja, J., 2021. Annual time series of global VIIRS nighttime lights derived from monthly averages. Remote Sensing 13 (5), 922.

Henderson, J.V., Storeygard, A., Weil, D.N., 2012. Measuring Economic Growth from Outer Space. Am. Econ. Rev. 102 (2), 994–1028.
HIFLD. Homeland Infrastructure Foundation-Level Data (HIFLD).
Huang, Q., Yang, X., Gao, B., Yang, Y., Zhao, Y., 2014. Application of DMSP/OLS Nighttime Light Images: A Meta-Analysis and a Systematic Literature Review. Remote Sensing 6 (8), 6844–6866.
Ji, W., Li, X., Chen, S., Ren, L., 2021. Transmission of SARS-CoV-2 via fomite, especially cold chain, should not be ignored. PNAS 118 (11) e2026093118.
Johns Hopkins Coronavirus Resource Center, *COVID-19 Data in Motion*. 2021.
Jia, P., et al., 2019. Earth observation: investigating non-communicable diseases from space. Ann. Rev. Publ. Health 40, 85–104.
Jia, P., 2019. Spatial lifecourse epidemiology. Lancet Planet. Health 3 (2), e57–e59.
Jia, P., 2020. Understanding the epidemic course in order to improve epidemic forecasting. GeoHealth 4 (10), e2020GH000303.
Jia, P., et al., 2020. Spatial lifecourse epidemiology reporting standards (ISLE-ReSt) statement. Health Place 61, 102243.
Klompmaker, J.O., Hart, J.E., Holland, I., Sabath, M.B., Wu, X., Laden, F., Dominici, F., James, P., 2021. County-level exposures to greenness and associations with COVID-19 incidence and mortality in the United States. Environ. Res. 199, 111331.
Kyba, C., Garz, S., Kuechly, H., de Miguel, A., Zamorano, J., Fischer, J., Hölker, F., 2015. High-Resolution Imagery of Earth at Night: New Sources Opportunities and Challenges. Remote Sensing 7 (1), 1–23.
Liu, Q., Sha, D., Liu, W., Houser, P., Zhang, L., Hou, R., Lan, H., Flynn, C., Lu, M., Hu, T., Yang, C., 2020. Spatiotemporal Patterns of COVID-19 Impact on Human Activities and Environment in Mainland China Using Nighttime Light and Air Quality Data. Remote Sensing 12 (10), 1576.
Lopez Bernal, J., Andrews, N., Gower, C., Gallagher, E., Simmons, R., Thelwall, S., Stowe, J., Tessier, E., Groves, N., Dabrera, G., Myers, R., Campbell, C.N.J., Amirthalingam, G., Edmunds, M., Zambon, M., Brown, K.E., Hopkins, S., Chand, M., Ramsay, M., 2021. Effectiveness of Covid-19 Vaccines against the B.1.617.2 (Delta) Variant. The New England J. medicine 385 (7), 585–594.
Ma, Y., et al., 2020. Effects of temperature variation and humidity on the death of COVID-19 in Wuhan China. Science The Total Environment 724, 138226.
Ma, T., Yin, Z., Zhou, A., 2018. Delineating Spatial Patterns in Human Settlements Using VIIRS Nighttime Light Data: A Watershed-Based Partition Approach. Remote Sensing 10 (3), 465.
Mayuri Chaturvedi, Ghosh, Tilottama, Laveesh Bhandari, 2011. *Assessing Income Distribution at the District Level for India Using Nighttime Satellite Imagery*. Proc. Asia-Pacific Adv. Netw. 32, 192–217.
Murphy, N., Boland, M., Bambury, N., Fitzgerald, M., Comerford, L., Dever, N., O'Sullivan, M.B., Petty-Saphon, N., Kiernan, R., Jensen, M., O'Connor, L., 2020. A large national outbreak of COVID-19 linked to air travel, Ireland, summer 2020. Eurosurveillance 25 (42).
Notari, A., 2021. Temperature dependence of COVID-19 transmission. Sci. Total Environ. 763, 144390.
Meng, Y., Zhu, V., Zhu, Y.e.a., 2021. *Co-distribution of Light at Night (LAN) and COVID-19 Incidence in the United States,*. PREPRINT (Version 2) available at Research Square.
Raifman, et al., *COVID-19 US State Policy Database.* 2020-09-15, Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor].
Small, C., Sousa, D., 2021. Spatiotemporal evolution of COVID-19 infection and detection within night light networks: comparative analysis of USA and China. Applied Network Science 6 (1), 10.
Sutton, P.C.E., Obremski, C.D.T., 2003. Building and Evaluating Models To Estimate Ambient Population Density. Photogramm. Eng. Remote Sens. 69 (5), 545–553.
Sutton, P., Roberts, D., Elvidge, C., Baugh, K., 2001. Census from Heaven: An estimate of the global human population using night-time satellite imagery. Int. J. Remote Sens. 22 (16), 3061–3076.
Shen, Y., et al., 2020. Community Outbreak Investigation of SARS-CoV-2 Transmission Among Bus Riders in Eastern China 180 (12), 1665–1671.
Sutton, P.C.E., C.D.; Ghosh, T., 2007. ***Estimation Of Gross Domestic Product At SubNational Scales Using Nighttime Satellite Imagery.*** International Journal of Ecological Economics and Statistics, **8**: p. 5-21.
The COVID Tracking Project, *The COVID Tracking Project*. 2021.
The United States Census Bureau, *American Community Survey Data*. 2021.
The U.S. Bureau of Economic Analysis. *Gross Domestic Product by State, 4th Quarter 2020 and Annual 2020*. 2021; Available from: https://www.bea.gov/news/2021/gross-domestic-product-state-4th-quarter-2020-and-annual-2020-preliminary#:~:text=Real%20gross%20domestic%20product%20(GDP,U.S.%20Bureau%20of%20Economic%20Analysis.
Wang, J., et al., 2021. Impact of temperature and relative humidity on the transmission of COVID-19: a modelling study in China and the United States. BMJ Open 11, e043863.
Wu, X., Nethery, R.C., Sabath, M.B., Braun, D., Dominici, F., 2020. *Air pollution and COVID-19 mortality in the United States: Strengths and limitations of an ecological regression analysis*. Science Advances 6 (45).
Xu, G., et al., 2021. *Lockdown induced night-time light dynamics during the COVID-19 epidemic in global megacities*. Int. J. Appl. Earth Observ. Geoinform. 102421.
Yau, K.K.W., Wang, K., Lee, A.H., 2003. Zero-Inflated Negative Binomial Mixed Regression Modeling of Over-Dispersed Count Data with Extra Zeros. Biometrical Journal 45 (4), 437–452.
Zhang, D., Jones, R.R., Powell-Wiley, T.M., Jia, P., James, P., Xiao, Q., 2020. A large prospective investigation of outdoor light at night and obesity in the NIH-AARP Diet and Health Study. Environ Health 19 (1).
Zhang, X., Mallick, H., Tang, Z., Zhang, L., Cui, X., Benson, A.K., Yi, N., 2017. Negative binomial mixed models for analyzing microbiome count data. BMC Bioinf. 18 (1).
Zhang, Q., Seto, K.C., 2011. Mapping urbanization dynamics at regional and global scales using multi-temporal DMSP/OLS nighttime light data. Remote Sens. Environ. 115 (9), 2320–2329.