

Co-distribution of light at night (LAN) and COVID-19 incidence in the United States

Yidan Meng

Yale University School of Public Health

Vincent Zhu

Yale University School of Public Health

Yong Zhu (✉ yong.zhu@yale.edu)

Yale University School of Public Health <https://orcid.org/0000-0002-6957-1114>

Research

Keywords: Light at night (LAN), COVID-19

Posted Date: February 22nd, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-195410/v1>

License:   This work is licensed under a Creative Commons Attribution 4.0 International License. [Read Full License](#)

Version of Record: A version of this preprint was published at BMC Public Health on August 4th, 2021. See the published version at <https://doi.org/10.1186/s12889-021-11500-6>.

Abstract

Background: Light at night (LAN) as a circadian disruption factor may affect the human immune system and consequently increase the susceptibility or severity of infectious diseases, including COVID-19. COVID-19 infections spread differently in states in the United States (US). The current analysis aimed to test whether there is an association between LAN and COVID-19 cases in 4 selected US states: Connecticut, New York State, California and Texas.

Methods: We analyzed clustering patterns of COVID-19 cases in ArcMap and performed the multiple linear regression model using data of LAN and COVID-19 incidence with adjustment for confounding variables including population density, percent below poverty, and racial factors.

Results: Hotspots of LAN and COVID-19 cases are located in large cities or metro-centers for all 4 states. LAN intensity is associated with cases/1k for overall, and lockdown durations in New York State, and Connecticut ($P < 0.001$), but not in Texas and California. The overall cases rates are significantly associated with LAN in New York State ($P < 0.001$) and Connecticut ($P < 0.001$).

Conclusions: We observed a significant positive correlation between LAN intensity and COVID-19 cases-rate/1k, suggesting that circadian disruption of ambient light may increase the COVID-19 infection rate possibly by affecting an individual's immune functions. Furthermore, differences in demographic structure and lockdown policies in different states play an important role in COVID-19 infections.

Introduction

Mammalian circadian rhythms, controlled by a neurological master clock located in the suprachiasmatic nucleus (SCN) and peripheral clocks of somatic cells, regulate a number of biological and physiological processes including the human immune system [1, 2]. Because immune responses play a major role in fighting against virus infections [3], disrupted circadian rhythm may adversely influence immune functions and consequently increase virus infectivity and ability to replicate inside the hosts [4-6].

Circadian disruptions can be caused by sleep deprivation, night shift work, frequent air traveling, circadian gene alterations and light at night (LAN) exposures [7-10]. Sleep deprivation has been associated with increased susceptibility to gut infection [11]. A higher incidence and severity of respiratory infections has been reported among night shift workers [12]. These findings support a significant relationship between disrupted circadian rhythms and increased individual's vulnerability to infectious diseases and suggest that excess risk could also be observed among individuals with high LAN exposures for the infection of COVID-19 [2], a coronavirus causing the global pandemic in 2020.

There are also studies analyzing LAN exposure and various cancer types, such as breast cancer, prostate cancer, thyroid cancer, and non-Hodgkin Lymphoma [7, 9, 13, 14]. Both global and regional studies have shown that there is a significant association between intensity of light at night and breast cancer [7, 14-16]. These findings suggest that LAN as a circadian disruption can influence immune system and hormone release and in turn affect the susceptibility of infectious diseases as well.

Light at night comes from either ambient light or indoor artificial light exposures. Excessive exposure to LAN may generate light pollution that may cause adverse effects on immune functions [17] and alter circadian gene functions in the SCN [13]. City-level LAN intensity can be measured by using the U.S. Defense Meteorological Satellite Program (DMSP). In the present analysis, we investigate whether exposure to LAN is associated with COVID-19 incidence in major cities in four selected US states: Connecticut, New York, Texas, and California, that represent different geological locations.

Methods

Data sources We obtained COVID-19 cases and tested data from local health departments. Specifically, we obtained COVID-19 data for Connecticut from the Connecticut State Department of Public Health (<https://data.ca.gov/dataset/covid-19-cases>), for New York State from the Open NY Program (<https://data.ny.gov/browse?tags=covid-19>), for Texas from Texas Health and Human Services (<https://dshs.texas.gov/coronavirus/additionaldata.aspx>), and for California from the California Open Data Portal (<https://data.ca.gov/dataset/covid-19-cases>). Data of COVID-19 cases were categorized in 3 groups: duration of overall period (to August 24th, 2020), period of lockdown, and period of reopening in each state according to local state policy on their government websites. In summary, we obtained 62 data points for New York State, 167 data points for Connecticut, 56 data points for California, and 254 data points for Texas, based on data availability for either county or town level.

LAN intensity data were extracted from satellite images of nighttime light intensity created by NASA Earth Observatory [18]. We collected demographic data, including factors of income and poverty, race/ethnicity, and population density. County level demographic data for California, Texas, and New York State were obtained from the US Census (<https://data.census.gov>), and town level demographic data for Connecticut were obtained from local public health departments (<https://portal.ct.gov/DPH>). County and town level boundaries data were from local transportation or planning departments (<http://gis.ny.gov/gisdata>; <https://data.ct.gov>; <https://gis-txdot.opendata.arcgis.com>; <https://gisdata-caltrans.opendata.arcgis.com>).

Geographic information system (GIS) mapping ArcMap (<https://desktop.arcgis.com/en/arcmap>) was used to generate visualized hotspot or density maps for COVID-19 case rate data and LAN data. The *Kernel Density* (KD) in the ArcMap 10.8.1 was used to calculate density from neighborhood features which are

closed to interested features, by using the kernel function to create a smooth raster layer from points or polylines. Search radius was calculated by spatial configuration and total number of points in the dataset, and equal breaks were used for symbology for Connecticut, Texas, and California. Natural breaks were used for symbology for New York State because the LAN level in New York City is much higher than other cities in the New York state and natural breaks can better present data with large differences of inherent groups. Because of the huge differences in LAN data of New York City compared to other cities in New York State, we specifically analyzed the spatial pattern of LAN and COVID-19 cases rate and performed the Geographically Weighted Regression (GWR). The Spatial Autocorrelation (Global Moran's I) tool was used to test spatial patterns (clustered, dispersed, or random) of points for cases/1k during lockdown, reopening, and overall, for New York State. GWR was used to understand regional variation of geo-data.

Statistically analysis To analyze the correlation between LAN intensities and COVID-19 cases rate per 1,000 people, we built multiple linear regression models with variables of nonwhite rate, percent below poverty, and population density from the US Census, 2015: ACS 5-Years Estimates Subject Tables (<https://data.census.gov>). Multiple regression models were performed using SAS 9.4 and RStudio.

Results

To represent different regions of the U.S., we included four states—Connecticut, New York State, California and Texas—in the study. Generally, hotspots of LAN data are located in large cities or metro-centers for all four states tested, such as Hartford, CT, New York City, NY, Dallas, TX, and San Francisco, CA.

The maximum LAN intensity calculated by ArcMap tool is 254.68, 254.68, 235.14 and 190.30 in New York State, Connecticut, Texas and California, respectively. The mean LAN intensity is 92.31 in Connecticut and around 40 in other three states. The maximum COVID-19 case-rate is highest in Texas among all four states, which is around 67 cases/1k people, followed by California (around 47 cases/1k people), New York State (around 43 cases/1k people), and lowest in Connecticut around 27 cases/1k people at the time of data collection.

For Connecticut and New York State, the hotspots of LAN intensity, COVID-19 cases rate/1k during lockdown, and COVID-19 cases rate/1k during overall durations shared similar patterns that clustered around major cities in the state (**Fig. 1**), but were slightly different after reopening. Differently, in Texas and California, there were inconsistent patterns of hotspots for the four interested variables, but they share the similar trend that the hotspot-areas were similar for reopening and overall durations.

During the lockdown period, the hotspots of COVID-19 cases rate shared similar patterns with LAN data in Connecticut and New York states, but have different geo-patterns in California and Texas. During the reopening period, the locations of hotspots of COVID-19 cases rate were very different from the LAN data map in NY state, Texas and California. The hotspots patterns of COVID-19 cases rate during the overall period are similar to the LAN map in Connecticut and New York. On the contrary, the overall hotspots of COVID-19 cases rate were similar to the reopening period in California and Texas.

Analysis of multivariable regression models also revealed similar patterns to the ones that geo-patterns showed. The cross comparison among four states showed that there were statistically significant correlations between LAN intensity and cases/1k for overall, and lockdown durations in New York State, and Connecticut ($P < 0.001$). There was no statistically significant association between LAN intensity and cases/1k, for overall, lockdown, or reopening durations, in Texas and California (**Table.1**). The overall cases rates were significantly associated with LAN in New York State ($p < 0.001$) and Connecticut ($p < 0.001$), that every 1 unit increase of LAN will have a 15.6% increase in the overall cases rate in New York State, and a 3.7% increase in Connecticut. During the lockdown period, there were similar results for the overall period in NY state ($p < 0.001$) and Connecticut ($p < 0.001$). During the reopening period, there was a significant small positive association between cases rate and LAN in Connecticut ($p < 0.001$). Based on R-squared results, the state-specific regression models could explain more variations in New York State ($R^2 = 0.80, 0.78$ and 0.40) and Connecticut ($R^2 = 0.58, 0.57$ and 0.21), compared with data in California ($R^2 = 0.22, 0.37, 0.23$) and Texas ($R^2 = 0.13, 0.02, 0.12$), for the overall, lockdown and reopening durations.

Discussion

Our findings demonstrate significant positive correlations between LAN intensity and COVID-19 cases rate per 1000 in two of the four US states (Connecticut and New York State) studied. These findings support the proposed hypothesis that high LAN exposures may disrupt circadian rhythms and consequently increase individual's vulnerability to infectious diseases including COVID-19 [2].

The co-distribution of light at night (LAN) and COVID-19 incidence was more evident in New York State and Connecticut. The regression models we built can explain more variations of COVID-19 data in New York State and Connecticut than that in Texas and California. This difference might be due to different lockdown/reopening policies in different states. For example, California had a shorter lockdown period, and there were no formal Stay at Home orders in Texas, compared to New York State which has similar total COVID-19 cases and population. The less strict policies may introduce more social factors into the analysis that we cannot determine easily at this point. We also found some inconsistent results of factors of social determinants, which might be due to different social and demographic structures in different states. Moreover, the different COVID-19 testing policy and availability in different states might introduce more variations into our analysis.

Data points for all four states are relatively less likely to have a very solid statistical analysis, but in order to discuss different policy impacts, we chose to analyze states specially because of the heterogeneity for each state. Data used in the study are at either town or county levels, which aggregate individual data into large spatial area levels and may introduce ecological fallacy. To reduce the effect of ecological bias, we performed GWR models for New York State. This approach improved the analysis and generated a very high level of R-squared from the regression models (**Supplement Table 1.**).

The current analysis focused on the COVID-19 test rate, and information on the severity of infected patients were not included due to its unavailability. Confounding factors considered in the analysis were percent below poverty, non-white rate and population density. More confounders, such as employment, education level, underlying health conditions, and proximity to healthcare facilities should also be considered in future studies if they are available.

Observations from our study are consistent with findings from a recent study that shows melatonin usage is significantly associated with a 28% reduced likelihood of a positive laboratory test result for COVID-19 [19]. Melatonin production in the pineal gland is sensitive to light and it has shown that even exposures of low intensity can suppress melatonin secretion [20]. LAN may reduce melatonin levels and consequently increase risk of COVID-19 infection.

Due to mental pressure, behavior, and daily routine changes during the pandemic, there are increasing concerns of sleep disturbances and circadian disruptions, especially for healthcare workers [5, 21, 22]. This provides more opportunities to analyze how circadian disruptions such as LAN correlate with hormone-related health outcomes and temporal immune dynamics [23].

In summary, both LAN intensity and COVID-19 cases rate are higher in major cities or metro-centers in all four states, due to the nature of cities of higher mobility, population density, etc. In the current study, we observed a significant positive correlation between LAN intensity and COVID-19 cases-rate/1000, which suggests that circadian disruption of ambient light may increase the COVID-19 infection rate possibly by affecting an individual's immune functions. Furthermore, differences of demographic structure and lockdown policies in different states play an important role in COVID-19 infections.

Declarations

Ethical Approval and Consent to participate

Not applicable

Consent for publication

Not applicable

Availability of supporting data

Not applicable

Competing interests

The authors declare no potential conflicts of interest.

Funding

This work was partially supported by funds from Yale University.

Authors' contributions

Acquisition of data: Yidan Meng

Analysis and interpretation: Yidan Meng, Yong Zhu

Writing, review, and/or revision of the manuscript: Yidan Meng, Vincent Zhu, Yong Zhu

Conception and design: Yong Zhu

Acknowledgements

Not applicable

References

1. Maiese K: **Moving to the Rhythm with Clock (Circadian) Genes, Autophagy, mTOR, and SIRT1 in Degenerative Disease and Cancer.** *Curr Neurovasc Res* 2017, **14**(3):299-304.

2. Silva FRD, Guerreiro RC, Andrade HA, Stieler E, Silva A, de Mello MT: **Does the compromised sleep and circadian disruption of night and shiftworkers make them highly vulnerable to 2019 coronavirus disease (COVID-19)?** *Chronobiol Int* 2020, **37**(5):607-617.
3. Rouse BT, Sehrawat S: **Immunity and immunopathology to viruses: what decides the outcome?** *Nat Rev Immunol* 2010, **10**(7):514-526.
4. Anderson G, Reiter RJ: **Melatonin: Roles in influenza, Covid-19, and other viral infections.** *Rev Med Virol* 2020, **30**(3):e2109-e2109.
5. Erren TC, Lewis P: **SARS-CoV-2/COVID-19 and physical distancing: risk for circadian rhythm dysregulation, advice to alleviate it, and natural experiment research opportunities.** *Chronobiol Int* 2020, **37**(7):1106-1109.
6. Meira E Cruz M, Miyazawa M, Gozal D: **Putative contributions of circadian clock and sleep in the context of SARS-CoV-2 infection.** *Eur Respir J* 2020, **55**(6):2001023.
7. Garcia-Saenz A, Sanchez de Miguel A, Espinosa A, Valentin A, Aragonés N, Llorca J, Amiano P, Martín Sanchez V, Guevara M, Capelo R *et al*: **Evaluating the Association between Artificial Light-at-Night Exposure and Breast and Prostate Cancer Risk in Spain (MCC-Spain Study).** *Environ Health Perspect* 2018, **126**(4):047011.
8. IARC: **Carcinogenicity of night shift work.** *Lancet Oncol* 2019;**20**(8):1058-1059 2019(1474-5488 (Electronic)).
9. Stevens RG, Zhu Y: **Electric light, particularly at night, disrupts human circadian rhythmicity: is that a problem?** *Philos Trans R Soc Lond B Biol Sci* 2015, **370**(1667).
10. Touitou Y, Reinberg A, Touitou D: **Association between light at night, melatonin secretion, sleep deprivation, and the internal clock: Health impacts and mechanisms of circadian disruption.** *Life Sci* 2017, **173**:94-106.
11. Anderson JR, Carroll I, Azcarate-Peril MA, Rochette AD, Heinberg LJ, Peat C, Steffen K, Manderino LM, Mitchell J, Gunstad J: **A preliminary examination of gut microbiota, sleep, and cognitive flexibility in healthy older adults.** *Sleep Med* 2017, **38**:104-107.
12. Archer SN, Oster H: **How sleep and wakefulness influence circadian rhythmicity: effects of insufficient and mistimed sleep on the animal and human transcriptome.** *J Sleep Res* 2015, **24**(5):476-493.
13. Stevens RG, Blask DE, Brainard GC, Hansen J, Lockley SW, Provencio I, Rea MS, Reinlib L: **Meeting report: the role of environmental lighting and circadian disruption in cancer and other diseases.** *Environ Health Perspect* 2007, **115**(9):1357-1362.
14. Zhong C, Franklin M, Wiemels J, McKean-Cowdin R, Chung NT, Benbow J, Wang SS, Lacey Jr JV, Longcore T: **Outdoor artificial light at night and risk of non-Hodgkin lymphoma among women in the California Teachers Study cohort.** *Cancer Epidemiology* 2020, **69**:101811.
15. Keshet-Sitton A, Or-Chen K, Huber E, Haim A: **Illuminating a Risk for Breast Cancer: A Preliminary Ecological Study on the Association Between Streetlight and Breast Cancer.** *Integr Cancer Ther* 2017, **16**(4):451-463.

16. Kloog I, Stevens RG, Haim A, Portnov BA: **Nighttime light level co-distributes with breast cancer incidence worldwide.** *Cancer Causes Control* 2010, **21**(12):2059-2068.
17. Nelson RJ: **Seasonal immune function and sickness responses.** *Trends Immunol* 2004, **25**(4):187-192.
18. **Global Map Downloads - 2016 Grayscale** [<https://earthobservatory.nasa.gov/features/NightLights>]
19. Zhou Y, Hou Y, Shen J, Mehra R, Kallianpur A, Culver DA, Gack MU, Farha S, Zein J, Comhair S *et al*: **A network medicine approach to investigation and population-based validation of disease manifestations and drug repurposing for COVID-19.** *PLoS Biol* 2020, **18**(11):e3000970.
20. Cajochen C, Munch M, Kobiacka S, Krauchi K, Steiner R, Oelhafen P, Orgul S, Wirz-Justice A: **High sensitivity of human melatonin, alertness, thermoregulation, and heart rate to short wavelength light.** *J Clin Endocrinol Metab* 2005, **90**(3):1311-1316.
21. Gao C, Scullin MK: **Sleep health early in the coronavirus disease 2019 (COVID-19) outbreak in the United States: integrating longitudinal, cross-sectional, and retrospective recall data.** *Sleep Med* 2020, **73**:1-10.
22. Morin CM, Carrier J, Bastien C, Godbout R, Canadian S, Circadian N: **Sleep and circadian rhythm in response to the COVID-19 pandemic.** *Can J Public Health* 2020, **111**(5):654-657.
23. Shirazi J, Donzanti MJ, Nelson KM, Zurakowski R, Fromen CA, Gleghorn JP: **Significant Unresolved Questions and Opportunities for Bioengineering in Understanding and Treating COVID-19 Disease Progression.** *Cell Mol Bioeng* 2020:1-26.

Tables

Table 1: Summary table of Regression models of association between COVID case-rate and LAN intensity with covariants: nonwhite rate, percent below poverty, and population density for overall, lockdown, and reopening period for New York State, Connecticut, California, and Texas.

Variables	N	Overall cases rate			Lockdown cases rate			Reopening cases rate		
		Beta	P-value	R ²	Beta	P-value	R ²	Beta	P-value	R ²
New York	62	0.156	6.61E-06	0.80	0.148	1.34E-05	0.78	0.003	0.4072	0.40
Connecticut	167	0.037	6.10E-11	0.58	0.030	8.90E-11	0.57	0.006	0.0080	0.21
Texas	254	0.043	0.1602	0.13	0.001	0.8682	0.02	0.041	0.1558	0.12
California	56	0.110	0.1428	0.22	0.001	0.8960	0.37	0.110	0.1297	0.23

Note:

N: county level data points for states of NY, CA, TX; town level data points for CT;

Beta: beta-estimate for the regression model, representing the coefficient of models;

R²: measured the goodness of fitting for regression models, representing percent of data that is able to be explained by the model.

Supplement Table 1: Spatial Autocorrelation test (Global Moran's I) of cases/1k per 1k during lockdown, reopening and overall, and Geographically Weighted Regression of cases /1k with variables: LAN2016, nonwhite rate, percent below poverty, and population density during lockdown, reopening and overall, for New York State.

	Lockdown cases rate	Reopening cases rate	Overall cases rate
Spatial Autocorrelation (Global Moran's I) (Spatial Statistics) for data in New York State			
Moran's Index	0.844	0.327	0.412
Variance	0.005	0.005	0.005
Z-score	11.81	4.71	6.11
P-value	<0.0001	<0.0001	<0.0001
Geographically Weighted Regression of COVID cases /1k in New York State			
Mean Coefficient (SD)	0.070 (0.026)	0.011 (0.0002)	0.129 (0.024)
Bandwidth	197257	609157	218270
Residual squares	562.9	14.91	997.09
Sigma	3.487	0.521	4.561
R ²	0.912	0.557	0.862
R ² adjusted	0.884	0.508	0.825

Note:

Moran's Index: The tendency of geo-clustering or geo-dispersion. A positive Moran's I show the tendency of geo-clustering;

Z-score: the critical value for test under standard normal distribution;

Bandwidth: distance band or neighbors used for each local regression equation;

Residual squares: sum of squared residuals, smaller the measure, the closer the fit of GWR models to observed data;

Sigma: square root of the normalized residual sum of squares represent standard deviation for residuals;

Figures

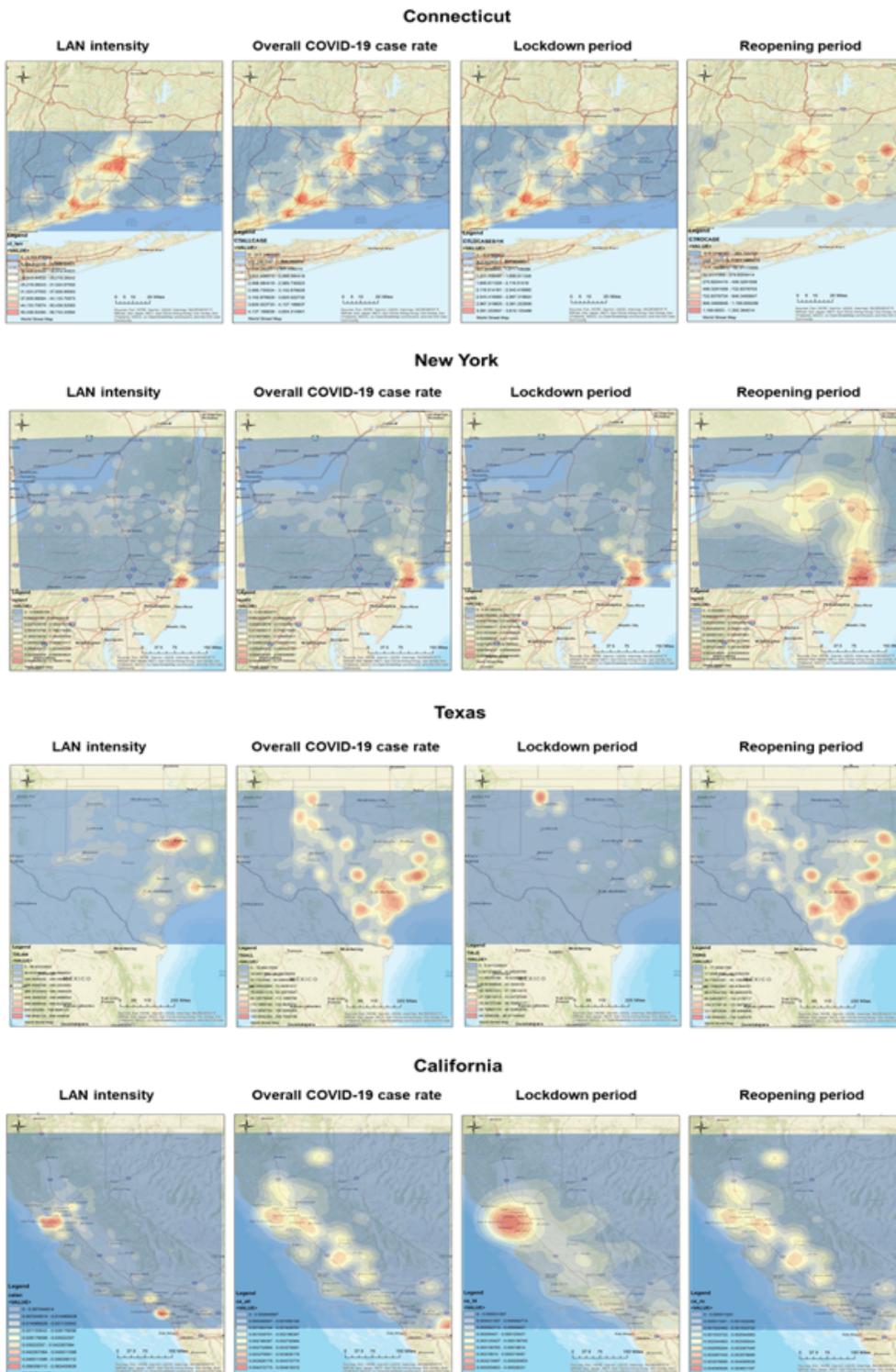


Figure 1

Co-distribution of light at night (LAN) and COVID-19 incidence in US states of Connecticut, New York, Texas and California. Kernel density maps of LAN intensity and COVID-19 cases rate for Connecticut, New York State, Texas, and California. There are significant positive correlations between LAN intensity and COVID-19 cases rate in Connecticut and New York State; no consistent correlation between LAN intensity and COVID-19 cases rate in Texas and California.

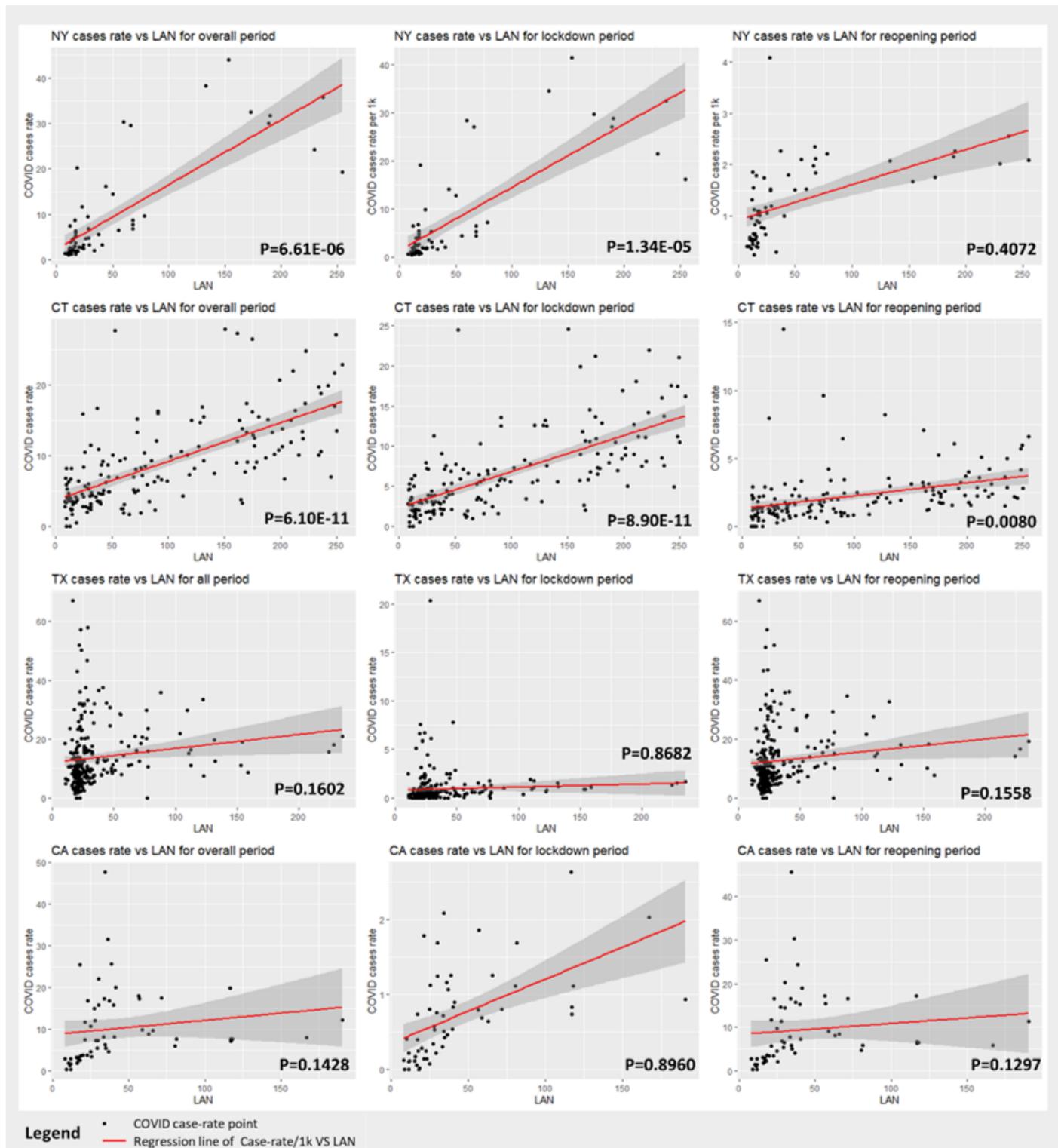


Figure 2

Scatter points of case-rate/1k, LAN intensity and regression lines in US states of Connecticut, New York, Texas and California. P values indicate significance of association between COVID case-rate and LAN intensity from Regression models.