Title:

Passive, open access data measures movement and predicts COVID-19 cases

Authors:

Christina L Faust[1]· Brian Lambert[1] · Cale Kochenour[1] · Anthony C. Robinson[2] · Nita Bharti[1]

Affiliations:

[1] Center for Infectious Disease Dynamics, Department of Biology, Eberly College of Science, Pennsylvania State University, University Park, PA, USA

[2] GeoVISTA Center, Department of Geography, College of Earth and Mineral Sciences, Pennsylvania State University, University Park, PA, USA

Correspondence to:

[CLF] czf134@psu.edu; [NB] nita@psu.edu

Keywords:

college towns; nighttime radiance; traffic volume; SARS-CoV-2

ABSTRACT

Background. In the months following the global spread of SARS-CoV-2, the lack of effective pharmaceutical interventions led to widespread implementation of behavioral interventions aimed at reducing contacts and transmission. In the US, state and local governments introduced and enforced the bulk of interventions, including university closures. As universities closed, student departures decreased the total population size of college towns while state-level interventions decreased contacts among remaining residents. Though the pandemic continues without pharmaceutical interventions, businesses have begun to reopen, and many universities have resumed operations. These actions have increased contacts and population sizes in college towns. Monitoring movement to implement adaptive policies will be critical for outbreak management.

Methods. We use publicly available remotely-sensed nighttime lights and traffic cameras to measure the impact of restriction policies on movement and activities in the university town of State College, and the surrounding areas of Centre County, Pennsylvania, USA.

Results. At the county level, nighttime radiance did not differ significantly across restriction phases and largely reflected seasonal fluctuations seen in previous years. Throughout the county, traffic volumes were lowest during the most severe period of restrictions ('Red phase' in Pennsylvania). As restrictions eased, traffic volumes grew, indicating increased movement within and between population centers. We show that real-time, publicly available traffic data captured behavioral responses and compliance to different restrictions phases. We also demonstrate that these increases in activity levels precede increases in reported COVID-19 cases.

Discussion. Passively collected data can measure population-level movement in response to restrictions and changes in these measured movements are reflected in observed SARS-CoV-2 transmission. Measuring these changes in movements and contacts in near real time can inform local adaptive interventions to curtail outbreaks.

MAIN TEXT

Introduction

Sudden Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2) was first detected in Wuhan, China, in December 2019.[1] Global spread by March 2020 resulted in a pandemic (Figure 1).[2] In humans, it causes the respiratory disease known as Coronavirus disease 2019, or COVID-19, and its clinical presentation includes a variety of symptoms, ranging from mild to severe or fatal.[3] There are currently no pharmaceutical interventions available for disease prevention or treatment. Outbreak management is focused on behavioral interventions to reduce transmission.[4,5]

Throughout the pandemic, behavioral intervention policies have varied widely between and within countries. China enforced severe interventions early, closing schools, workplaces, roads, and public transit, and requiring mandatory quarantine of uninfected people across entire provinces.[6] These reductions in contacts between individuals were effective and interrupted transmission in Hubei Province.[7] Other countries have imposed less strict lockdowns, with many relying on recommendations instead of legally enforceable restrictions.

Most universities in the United States cancelled face-to-face instruction mid-semester in the spring of 2020. Closures sent many students to family homes, emptying campuses and reducing population size and density in college towns. However, almost half of US universities have returned or will return to some degree of in-person teaching in the fall of 2020, with many more planning a return in the spring of 2021.[8] University reopenings increase population sizes as students return to campus, and are compounded with increased mixing as activities resume. College campuses are fundamentally designed to facilitate interactions, which must now be deliberately and drastically curtailed to prevent SARS-CoV-2 transmission.[9] Some universities rapidly pivoted to increase online engagement in response to large numbers of reported cases shortly after reopening. Initial modelling suggests that effective outbreak management on college campuses will require high compliance and widespread adoption of behavioral restrictions.[10] Testing capacity is still insufficient in most settings and delayed results prevent effective contract tracing.[11] It is essential for universities and surrounding communities to be able to measure the increases in population size and changes in human movement patterns as campuses reopen. Measuring movement and contacts, which precede transmission by about two weeks[3,12], helps create early warnings, rapid adaptive control strategies, and preventative public health messaging to curb outbreaks.

There are many ways to measure human populations and movement to inform disease transmission. Data resolution varies across spatiotemporal scales, from targeted individual surveys and censuses to the passive surveillance of satellites and mobile phone usage data. Epidemiological efforts to estimate human movement and contacts have included tracking currency,[13] commercial air traffic to model long distance flows,[14] anthropogenic illumination to quantify seasonal population changes,[15] and mobile phones for individual mobility traces.[4] Privately owned mobile device data are anonymized and confined to national boundaries[16] and due to their confidential nature, can't be shared with policy makers. Aggregate device data obtained through third party vendors can be expensive, rely on opaque

algorithms, and are proprietary. To overcome limitations on data sharing and increase replicability, we use publicly available passively collected satellite data and traffic cameras to measure indicators of human movement.

State College, in Centre County, is home to The Pennsylvania State University's University Park campus (referred to as Penn State and PSU). It is the largest campus of the state's largest public institute of higher education. In Pennsylvania, policies to minimize movement and local contacts were first implemented at the county level in late March (Figure 1, Table 1). Compared to previous years and throughout progression through three restriction phases (Table 1), we assessed changes in population level characteristics using remotely-sensed nighttime lights. To measure movement across restriction phases, we applied machine learning algorithms to ongoing traffic camera data collection.

We demonstrate the use of open access data to monitor near-real-time behavioral changes related to institutional and governmental COVID-19 response policies and outbreak trajectories through July 2020. Our methods are broadly applicable across university towns and the approaches presented here can be adapted to inform policies and messaging going forward.

Methods

Study area

Centre County is located in central Pennsylvania and has an estimated population of 162,000 (Fig 1).[17] Approximately 40,000 of these residents are undergraduate students enrolled at Penn State.[18] The majority of undergraduates leave during school holidays, resulting in seasonal population fluctuations.

Study period

We used radiance data, traffic cameras, and epidemiological records to measure activity and responses to restriction policies focusing on the period from February 14 to July 2, 2020 (Figure 1). We divided our surveillance efforts into six temporal phases, which align with events and policies that were expected to change movement and behavior at the University and across the county (Table 1, Fig 1):

1) **Baseline** (Feb 14 – March 5): before restrictions were in effect and while undergraduate students were on campus
2) **Population Decline** (March 6 – March 17): before local restrictions were in effect, after students left campus for spring break, and encompassing March 16th transition to online instruction
3) **Local Restrictions** (March 18 -March 27): Mandated closure of all non-essential businesses (no county-wide restrictions)
4) **County Red** (March 28 -May 7): red phase restrictions (Table 1)
5) **County Yellow** (May 8 to May 28): yellow phase restrictions
6) **County Green** (May 29 to July 2): green phase restrictions

Radiance data covered the study period from Feb 14 to July 3, traffic data collection spans April 26-July 2 2020, and epidemiological records included Feb 14 to July 16.

Radiance Data

We acquired nighttime radiance data from NASA's Black Marble calibrated product.[19] Black Marble quantifies daily anthropogenic nighttime radiance at 500-meter spatial resolution. The radiance data are captured by the day night band (DNB) of the Visible Infrared Imaging Radiometer Suite (VIIRS). They have been collected and publicly available since 2012.[19] These data reveal information on changes in population-level distribution, abundance, movement, behavior, and economic activity.[19] Here, we use images from February 14 to July 2 of the pandemic year (2020, Fig. 1) and compare them to typical seasonal variation using images collected across the same academic calendars from 2016 to 2019 (Table S1).[20] We downloaded NASA VNP46A1 Daily At-Sensor Top of Atmosphere (TOA) Nighttime Radiance products as Hierarchical Data Format 5 (HDF5) files for central Pennsylvania. We extracted radiance values from each HDF5 file and saved the extracted values to GeoTiff files at a resolution of 351 m x 463 m. Only cloud-free pixels were used in the radiance analyses and pixels were masked using Black Marble's cloud cover values. Acquisition and pre-processing operations were completed in Python version 3·7·6 (SI).[21]

Radiance rasters were cropped to Centre County boundaries to match the spatial extent of epidemiological and traffic data (Fig 2A). Gridded population density (1 km^2) estimates for 2020[22] were used to identify pixels with fewer than 5 people/ km^2 (i.e. state forests, golf courses; Fig 2A; S1). The remainder of the county was divided into population quantiles that each represented approximately 20% of the population (Table S2). Images contaminated by lunar radiance during a full moon, one day prior and three days following were removed (Fig S2).

We calculated mean radiance values for each of the six temporal phases and each population quantile from 2016-2020. PSU events (i.e. Spring Break) do not occur on the same calendar dates each year, so the dates were adjusted for each year to overlap with corresponding academic periods and cover the same number of days in previous years (Table S1). Previous years were included to quantify typical fluctuations in radiance driven by fluxes in the resident student population and seasonal influences of snow albedo during winter and occlusion from emergent vegetation in spring. The distributions of radiance values were visualized as density plots to illustrate the frequency and range of radiance values observed in each phase, population density, and year.

Traffic Cameras

We collected data from 19 traffic cameras across Centre County to quantify numbers of vehicles on roads beginning on April 27, 2020 (Fig 3A). Twelve of these cameras were on interstates, state highways, or other roads that link towns in Centre County, which we refer to as 'connector' roads. The remaining seven cameras are on 'internal' roads for travel within towns. Cameras produce 24-hour live streams accessible online but are not archived. We captured and stored images from these live streams every 20 seconds. Using the Python package *cvlib* as a high-level interface to OpenCV[23,24] and

Google's open source TensorFlow software stack,[25] we identified and counted vehicles in each image. Parked vehicles were excluded. The number of vehicles captured by each camera was standardized by the images captured within that hour (SI).

We fit a series of generalized additive models (GAMs) to these standardized hourly counts. The effect of hour of the day in local time was modelled as a cyclic cubic regression spline to account for the continuity of the variable so the same smoother operates at the 12:00 am 'start' of a day and the 11:00 pm 'end' of a day (Fig S3). We included the following as predictor variables, with interactions between variables: day of week, weekend (binary variable), restriction phase, camera identity, number of lanes visible in camera image, and road type (internal or connector road). Multivariate GAMs were implemented in the package *mgcv* in R version 3·6·2.[26,27] All cameras had some infrequent and short duration gaps in image acquisition (Fig S4). Missing hourly data were predicted using the best fit GAM. The combined data (both observed and predicted) were used to estimate vehicle traffic volumes.

COVID-19 diagnostic testing results

We acquired total daily, county-level confirmed cases of COVID-19 in Centre County and the surrounding counties, as reported by the Pennsylvania Department of Health.[28] Cases are confirmed using the CDC's diagnostic tests to detect active infections using reverse transcriptase polymerase chain reaction (RT-PCR). When linking epidemiological data to movement data, we estimated a two-week lag between a transmission event and a reported case. This incorporates a 5-8 day[12] incubation period and 5·5-10·93 day[3] delay between the onset of symptoms to case confirmation. The case data included here extend to July 16, 2020, two weeks after movement data collection ends.

Results

Radiance fluctuations throughout the pandemic period compared to previous years

During 2020, the majority (86·18%) of the study area has low population density (5-99 people km^{-2}), which accounts for 21·98% of the total population in Centre County (Fig 2A, Table S2). Low population density areas have low radiance values (<2 nW cm^{-2} sr^{-1}) across all years and little variation between phases (Fig 2B).

The highest density areas (>2000 people km^{-2}) encompass a similar percentage of the total population (20·97%) in a much smaller percentage of the study area (0·69%). In these highest density areas, nighttime radiance was more variable between years and across phases of all years, not just 2020. In Baseline and during the Population Decline phase (Feb 14-March 17), we observed the greatest variation in radiance values, which we expect to relate to snowfall in these respective years (Table S4). Previous years included a proportion of pixels with high radiance values, above 100 nW cm^{-2} sr^{-1}. During the red, yellow, and green restriction phases, we observed lower radiance values across all years, less variation between years, and little variation between phases within years.

Although we expected a reduction in radiance in 2020, particularly during local and county restriction phases, we did not observe this. While 2020 was darker in the earlier phases compared to 2016, 2018 and 2019, the mean and variation of radiance was not significantly different from 2017 or 2018. During restriction phases, 2020 radiance values were indistinguishable from all other years, suggesting no detectable changes in nighttime lights in response to policies.

Changes in traffic patterns through restriction phases

Data collection on traffic patterns began during the red phase, from April 27, 2020 forward. Throughout restriction phases observed in this study, daily traffic volume followed a bell-shaped pattern, with troughs at 02:00 h and peaks between 12:00 and 18:00 h (Fig. 3B, Fig S3). The best-fit general additive model included splines fit to hourly average counts per camera, a binary weekend predictor, intervention phase, and an interaction between phase and road type (connector or local). This model explained 87·9% of variation observed in vehicle traffic.

Traffic volume significantly increased in each subsequent phase of easing behavioral restrictions, with greater increases during weekdays and on connector roads (Fig 3B,C). During the red phase, mean daily weekday traffic totals were 10,947 vehicles on internal roads and 7,409 vehicles on connector roads. After movement to the yellow phase, mean daily weekday traffic totals increased by 23% on internal roads (increased to 13,265 vehicles) and 31% on connector roads (increase to 9,709 vehicles). While there was only a small increase from yellow to green on internal roads (3% increase), there was a 22% increase in vehicles on connector roads during this transition.

During the red phase (March 28-May 7), the governor announced on May 2 that Centre County would be shifting to yellow on March 8, six days prior to the actual lifting of restrictions. Traffic increased when restrictions officially eased, rather than when the announcement was made on March 2, suggesting behaviors followed governmental policies during this observation period.

Spatial and temporal patterns of confirmed cases in PA counties

Through July 17, Centre County confirmed a total of 208 cases, which is 166·5 cumulative confirmed cases per 100,000 census population. Increases in case counts were not synchronous with surrounding counties (Fig 1). Different counties within the state also experienced vastly different levels of morbidity. By July 17, the most severely impacted county in PA, Philadelphia County, reported 1,468·5 confirmed cases per 100,000 people.[28] The epidemic is ongoing, with a surge in cases in western Pennsylvania, where Allegheny county (location of Pittsburgh) doubled the cumulative cases in two weeks prior to July 17 to 502·1 per 100,000 residents.

Epidemiological trends in relation to movement measured through traffic and radiance fluctuations

Transmission precedes official case reporting by approximately two weeks (see Methods[3,12]). Although there is a decrease in the seven-day moving average of daily radiance in higher population density (≥100 km 2) from Baseline to the Red phase, this is consistent with typical seasonal patterns observed

in previous years during the transition from winter to spring (Fig 2B, 4A). There is no discernable pattern in low population areas where the mean radiance is much lower.

When focusing on the ten weeks of traffic data, there is an observed increase in daily traffic volume from red to green phases, with significant increases in vehicle volume observed on all roads and the greatest increases on connector roads.

The daily COVID-19 case totals are low in Centre County throughout the study period. Cases initially increase during the population decline and local restriction phases. Although stochastic, red and yellow phases correlate with a downward trajectory or plateau in daily incidence. In the green phase, larger increases in road traffic precede increases in incidence in Centre County and the highest 7-day averages in this period (Fig 4B, C).

Discussion

Behavioral interventions are currently the primary public health tool for reducing SARS-CoV-2 transmission. We used publicly available remotely-sensed nighttime lights and traffic camera data to measure behavioral responses to state- and local- COVID-19 restrictions in a University town and surrounding areas. We use traffic data to measure changes in movement and compliance with policies. This data source can be used to monitor movement and adapt mitigation strategies during the fall semester. Centre County has not yet experienced substantial community spread and monitoring movement levels throughout the region will be important for preventing sustained transmission following early increases in movement and populations.

In the United States, approximately half of universities are planning some level of fall 2020 return to campus operations,[8] with many more resuming in the spring of 2021. University towns face three challenges with a return to operations: 1) an initial surge in susceptible (and possibly infected) individuals in university towns, as students and employees arrive on campus from different locations, many of which may have higher incidence of COVID-19 than the relatively empty college towns to which they will return, 2) rapid increases in mixing in formal educational and informal social settings between people who are returning from many different locations and local permanent residents, and 3) higher daily contact rates between and among students and employees as academic and social activities resume in some form. University towns are experiencing simultaneous increases in local population size and numbers of contacts, raising the likelihood of introduced cases and local transmission, respectively. Quantifying population fluxes and changes in movement and contacts will be key for developing adaptive policies and preventing ongoing transmission in these settings.

Passively collected traffic camera images measured movement in response to phased restriction policies. Traffic volume increased, particularly on connector roads, as restrictions were loosened. Local road traffic increased, but not as much as connector roads, which may be due to the function of each road type. In Centre County, local roads provide access to essential businesses that remained open throughout the restrictions, whereas connector roads are used for travel between places and may

reflect a return to business, childcare and in-person work. These results suggest that Centre County residents largely complied with county-level restrictions. However, a caveat is that this region has not yet transitioned from a less restrictive phase to a more restrictive phase so compliance with this cannot be assessed. The increase in movement, as measured by vehicle volume, precede an overall increase in incidence in COVID-19 cases.

Although traffic camera data were able to detect changes in movement from vehicle volume, nighttime radiance data did not show changes in response to population declines or pandemic restriction policies. While there was a decline in radiance from Baseline to Red Phase Restrictions in 2020, this a typical seasonal trend seen in all other years from 2016-2019. The higher radiance values in the early phases are likely driven by albedo affects from snow, which is also highly variable between years (Table S4). Although we expected to see decreases in nighttime radiance values during restriction phases, reflecting the closing of non-essential businesses and stay-at-home mandates, this was not detected for high or low population density areas. It is possible that the rural, dimly lit area made these changes difficult to detect. Changes in population sizes in this county are concentrated in a very small minority of high population density pixels, which would be difficult to detect in radiance values across this large area. We find that highly localized population changes do not have a county-level impact on radiance values. These findings are in contrast to large scale changes detected in China, but over an area with much higher population density.[29]

In Pennsylvania, restrictions were implemented across the state within a few days at the end of March. A surge in cases in Philadelphia, the most populous county in the state, located 300 km southeast of Centre County, catalyzed the statewide lockdown. Only 2 cases had been confirmed in Centre County when the Red Phase of restrictions was implemented. Transmission of cases decreased during the red phase and was lowest, as a daily average, during the yellow phase. The increase in movement during the green phase, as observed in traffic data, particularly on connector roads, coincided with a second wave of infections which was increasing through July 16.

Testing capacity increased in the later phases of this study, but the bimodal pattern of infections most likely demonstrates increased transmission, which was preceded by increased movement, rather than simply an increase in testing. Increases in movement following easing restrictions, particularly the transition from red to yellow, increased contacts and led to the uptick in cases during the green phase in Centre County.

The local hospital is a relatively small care facility with limited COVID-19 patient care capacity[28] and it will quickly become overwhelmed if cases surge, particularly with the influx of over 40,000 students. Quantifying changes in traffic, which correlate with transmission and precede increases in reported cases by about two weeks, provides an actionable metric for proactive policy and behavioral restrictions. This approach will be an important part of a larger plan for staying ahead of COVID-19 in college towns.[10]

The epidemic in the United States has been managed by local governments implementing policies at state and county levels. COVID-19 has spread widely through the nation's big cities and small towns. Moving forward, it will be important to monitor local situations and implement locally responsive interventions. In addition to improving testing capacity and reducing delays in returning results, measuring local population movement and size at multiple scales offers a useful strategy to measure behavior and adapt policies and messaging targeting case prevention. University towns will face a large increase in populations and contacts whenever they resume campus operations, which will inevitably increase local transmission of COVID-19. Monitoring movements and contacts through integrated university and community efforts will be necessary to avoid overwhelming local health care capacities.

Contributors

NB and AR designed the study. BL and CK contributed to data collection. CLF BL and NB analyzed the data and interpreted the results. All authors contributed to writing and approved the final version of the report.

Declaration of interests

All authors declare no competing interests.

Data sharing

All data used for this study are available from publicly available sources and are attributed in the methods section. Code to generate models and figures are available at: https://github.com/cfaustus/centre_co_movement

Acknowledgements

We would like to thank the Pennsylvania Department of Transportation, the Pennsylvania Department of Health, and NASA for making data publicly available in real-time. We also acknowledge financial support from Penn State University's Huck Institutes of the Life Sciences and the Institute for Computational and Data Sciences. The funding sources had no role in the design, execution or interpretation of the results presented here.

TABLES

Table 1. Pennsylvania COVID-19 Restriction Phases

Restriction Phase	Length of phase in Centre County	Work & congregate settings	Social settings
Red	40 days (March 28 – May 7)	• Life-sustaining businesses only • Schools closed for in-person teaching • Most childcare closed	• Stay at home order • Large gatherings prohibited • Travel limited to life-sustaining purposes • Restaurants/ bars limited to carry-out and delivery
Yellow	20 days (May 8 – May 28)	• Telework where feasible • Businesses with in-person operations must follow safety guidelines • Schools closed for in-person teaching • Some childcare open with safety orders	• Full stay at home restrictions lifted • In person retail allowed, but curbside/delivery preferred
Green	35+ days (May 29 – July 3, ongoing)	• Businesses reopen with CDC and PA DOH guidelines (reduced capacity, distancing, mask wearing)	• Gyms, spas, entertainment reopen pending guidelines • Aggressive mitigation orders lifted (individuals must follow CDC and PA DOH guidelines)

FIGURES

Figure 1. Timeline of policy interventions and cases in Centre County and adjacent counties. The temporal phases of this study are shown at the bottom of the figure, above the timeline. These indicate pre-restriction and restriction phases in Centre County. Phase transitions for surrounding counties are indicated underneath the respective epidemic curves. Cases are shown as the date of report (rather than the estimated transmission date as in Fig 4C). Centre County was the first county in the region to go under stay-at-home orders in the red phase on March 27 (Table 1). The north-central counties transitioned to slightly more relaxed yellow phase activities starting May 9 and green phases starting May 29, with transition dates varying slightly between counties.

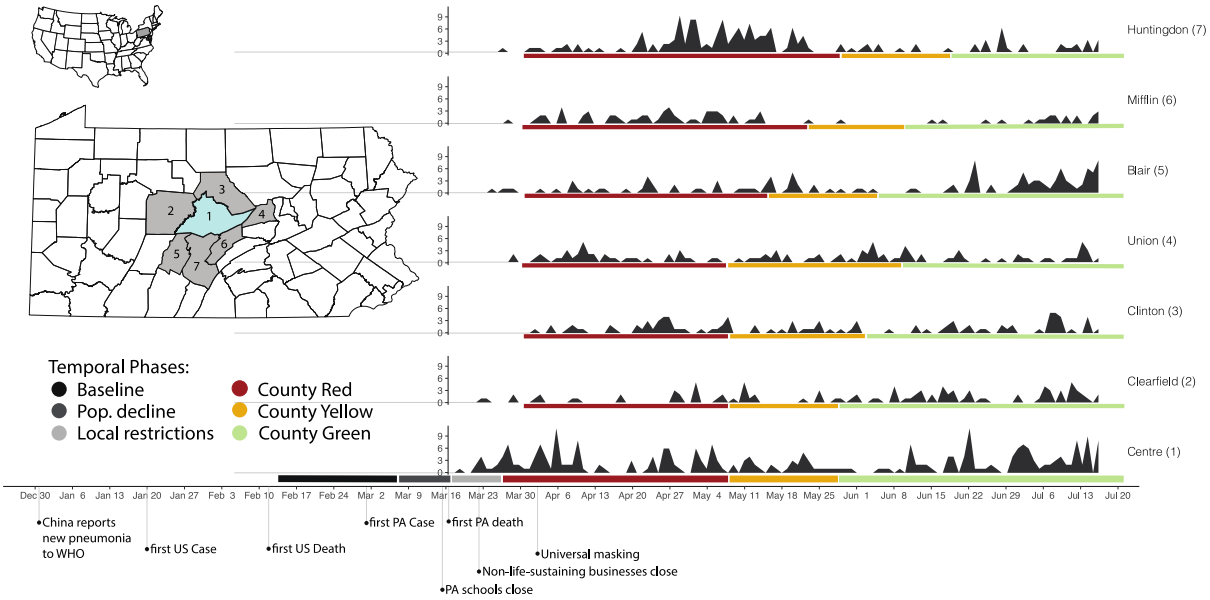


Figure 2. Changes in anthropogenic nighttime lights in 2020 compared to 2016-2019. (A) Centre County is a largely rural area. Locations with occupancy estimates below 5 people/km2 were excluded from the nighttime lights analysis (represents only 1·8% of the total county population). The remainder of the county was split into five population densities, which each accounted for approximately ⅕ of the population (Table S2, legend colors proportional to population sizes). State College and Bellefonte, the two largest towns in the county, are indicated as open grey circles for reference. (B) A density plot of mean pixel radiance for each pixel across each temporal phase. Each year from 2016-2020 shown separately, phase (rows), and population densities (columns). Note the change in x- and y- axis scales for each population density. Early phases showed more variation and higher radiance values in previous years. Overall, 2020 radiance values were not significantly different from the previous four years, particularly in the restriction phases.

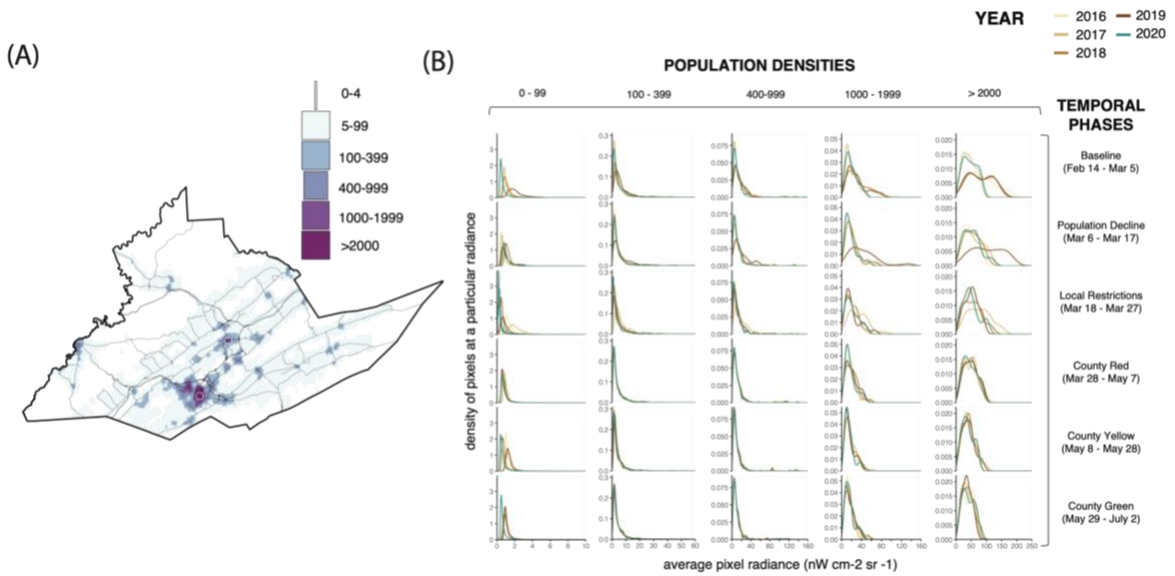


Figure 3. Traffic camera locations and volume throughout containment phases. (A) The internal traffic cameras were concentrated in State College and surrounding boroughs, whereas the connector roads covered a larger spatial extent. The direction of the arrow points in the direction in which the image was taken. (B) Traffic during all phases peaked between 12:00 and 18:00 h EST. Volume was significantly higher in the green phase with the greatest increases on connector roads. Typical rush hours (07:00-09:00 and 16:00-18:00) are shaded in grey. (C) Traffic increased in each phase with loosening of restrictions. The size of the point indicates the number of days of data collection. The lowest sample size was during weekends in the red phase (2 days) and the greatest number of observations was during weekdays in the green phase (26 days).

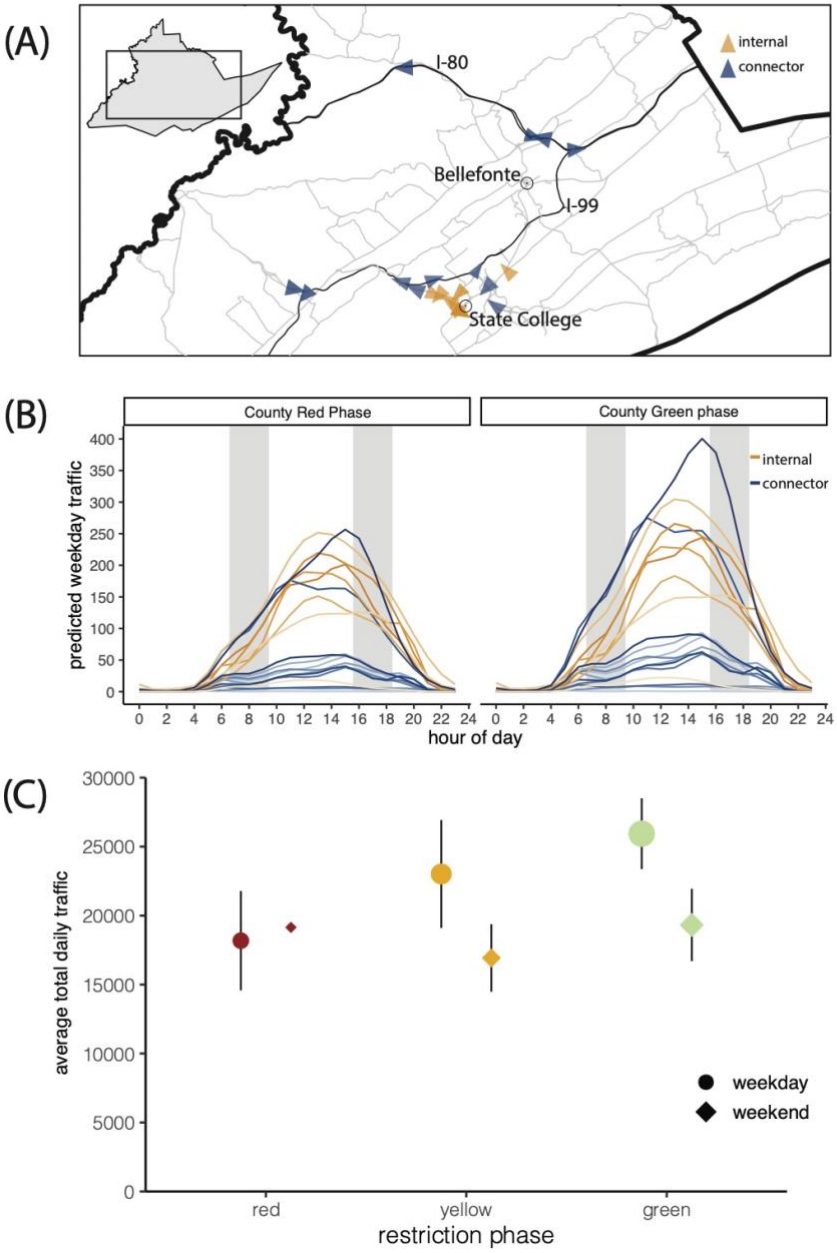
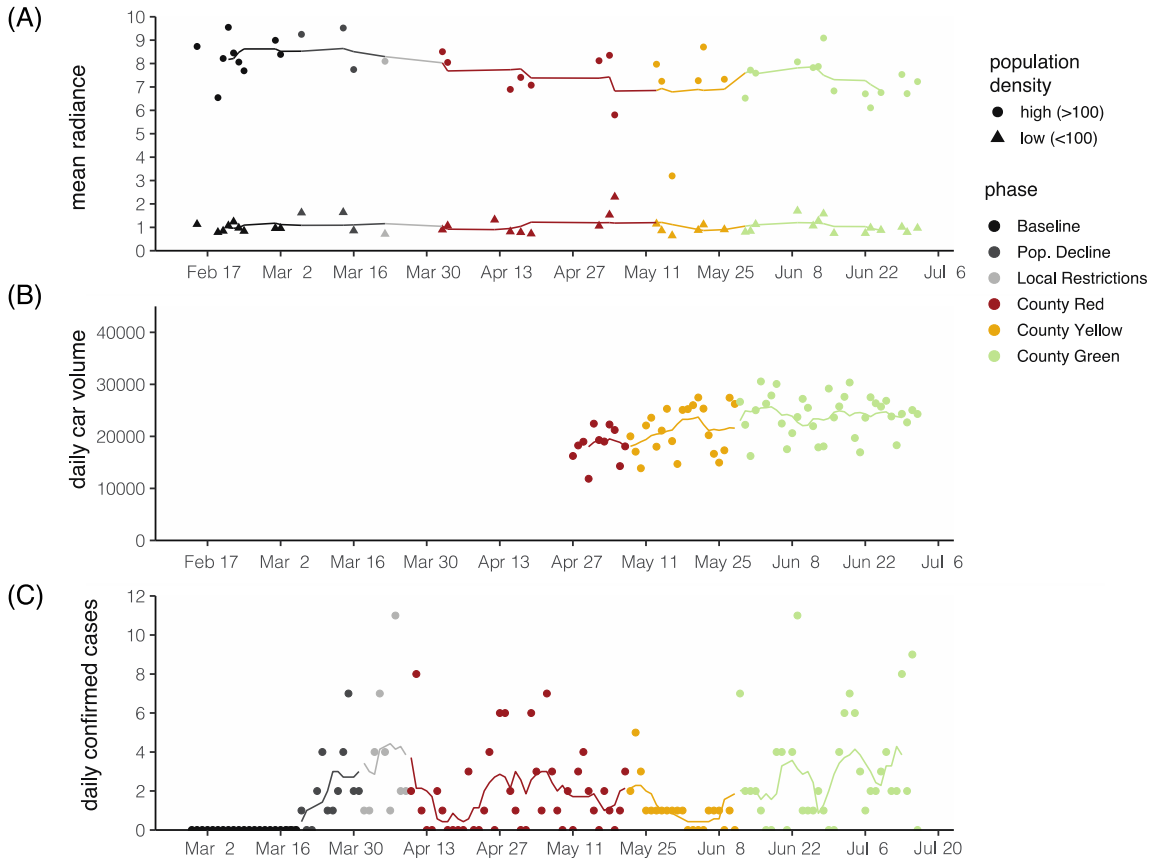


Figure 4. Case transmission aligns with movement across phases. (A) Mean radiance values are shown for days where nighttime images had a minimum of 70% coverage across the study area. (B) Traffic volume increases as restrictions are loosened. (C) Cases are shown by the date of confirmation, but aligned with radiance and traffic data (A,B) corresponding to the date of expected transmission, which we assume is 14 days prior to reporting date. Although confirmed cases are low, after an initial increase in cases in March, cases declined through red and yellow phases but showed an increase in green phase. Solid lines are rolling means from the previous 7 days.



References

1. Zhu N, Zhang D, Wang W, et al. A novel coronavirus from patients with pneumonia in China, 2019. *New England Journal of Medicine* 2020.
2. World Health O. WHO Director-General's opening remarks at the media briefing on COVID-19, 2020.
3. Li R, Pei S, Chen B, et al. Substantial undocumented infection facilitates the rapid dissemination of novel coronavirus (SARS-CoV-2). *Science* 2020; **368**(6490): 489-93.
4. Kraemer MU, Yang C-H, Gutierrez B, et al. The effect of human mobility and control measures on the COVID-19 epidemic in China. *Science* 2020; **368**(6490): 493-7.
5. Hatchett RJ, Mecher CE, Lipsitch M. Public health interventions and epidemic intensity during the 1918 influenza pandemic. *Proceedings of the National Academy of Sciences* 2007; **104**(18): 7582-7.
6. Kupferschimdt K, Cohen J. China's aggressive measures have slowed the coronavirus. They may not work in other countries. *Science* 2020.
7. Zhang J, Litvinova M, Liang Y, et al. Changes in contact patterns shape the dynamics of the COVID-19 outbreak in China. *Science* 2020; **368**(6498): 1481.
8. College Crisis Initiative at Davidson College, The Chronicle of Higher Education. C2i Dashboard for Reopening Plans. 2020. https://collegecrisis.shinyapps.io/dashboard/ (accessed August, 21, 2020)
9. Harris JE. Data From the COVID-19 Epidemic in Florida Suggest That Younger Cohorts Have Been Transmitting Their Infections to Less Socially Mobile Older Adults. *medRxiv* 2020: 2020.06.30.20143842.
10. Drake JM. Scenario analysis for an outbreak of COVID-19 on a university campus. Center for Ecology of Infectious Diseases: University of Georgia, 2020.
11. Kretzschmar ME, Rozhnova G, Bootsma MCJ, van Boven M, van de Wijgert JHHM, Bonten MJM. Impact of delays on effectiveness of contact tracing strategies for COVID-19: a modelling study. *The Lancet Public Health* 2020; **5**(8): e452-e9.
12. Backer JA, Klinkenberg D, Wallinga J. Incubation period of 2019 novel coronavirus (2019-nCoV) infections among travellers from Wuhan, China, 20–28 January 2020. *Eurosurveillance* 2020; **25**(5): 2000062.
13. Brockmann D, Hufnagel L, Geisel T. The scaling laws of human travel. *Nature* 2006; **439**(7075): 462-5.
14. Colizza V, Barrat A, Barthélemy M, Vespignani A. The role of the airline transportation network in the prediction and predictability of global epidemics. *Proceedings of the National Academy of Sciences* 2006; **103**(7): 2015-20.
15. Bharti N, Tatem AJ, Ferrari MJ, Grais RF, Djibo A, Grenfell BT. Explaining seasonal fluctuations of measles in Niger using nighttime lights imagery. *Science* 2011; **334**(6061): 1424-7.
16. Gonzalez MC, Hidalgo CA, Barabasi A-L. Understanding individual human mobility patterns. *Nature* 2008; **453**(7196): 779-82.
17. United States Census Bureau. QuickFacts: Centre County, Pennsylvania, 2019.
18. Penn State University. Penn State Undergraduate Admissions. 2020. https://admissions.psu.edu/apply/statistics/ (accessed 01 July 2020)
19. Miller SD, Mills SP, Elvidge CD, Lindsey DT, Lee TF, Hawkins JD. Suomi satellite brings to light a unique frontier of nighttime environmental sensing capabilities. *Proceedings of the National Academy of Sciences* 2012; **109**(39): 15706.
20. Levin N. The impact of seasonal changes on observed nighttime brightness from 2014 to 2015 monthly VIIRS DNB composites. *Remote Sensing of Environment* 2017; **193**: 150-64.
21. Python Software F. Python Language Reference, version 3.7.6. 2019.
22. WorldPop (School of Geography and Environmental Science UoSDoGaG, University of Louisville; Departement de Geographie, Universite de Namur) and Center for International Earth Science Information

Network (CIESIN), Columbia University. Global High Resolution Population Denominators Project - Funded by The Bill and Melinda Gates Foundation (OPP1134076). 2018. www.worldpop.org. (accessed 15 June 2020)

23. Ponnusamy A. cvlib - high level Computer Vision library for Python. 2018.

24. Bradski G. The OpenCV Library. Dr. Dobb's Journal of Software Tools; 2000.

25. Abadi M, Agarwal A, Barham P, et al. TensorFlow: Large-Scale Machine Learning on Heterogeneous Distributed Systems. *arXiv* 2016.

26. Wood S. Fast stable restricted maximum likelihood and marginal likelihood estimation of semiparametric generalized linear models. *Journal of the Royal Statistical Society (B)* 2011; **73**(1): 3-36.

27. R Core Team. R: A Language and Environment for Statistical Computing. Vienna, Austria: R Foundation for Statistical Computing,; 2019.

28. Pennsylvania National Electronic Disease Surveillance S. COVID-19 Data for Pennsylvania. https://www.health.pa.gov/topics/disease/coronavirus/Pages/Cases.aspx

29. Liu K, Ai S, Song S, et al. Population Movement, City Closure in Wuhan, and Geographical Expansion of the COVID-19 Infection in China in January 2020. *Clinical Infectious Diseases* 2020.

SUPPLEMENTARY INFORMATION

Passive, open access data measures movement and predicts COVID-19 cases

Supplementary Methods

I. Detailed timeline of events

II. Black Marble Pre-Processing Details

III. Details of traffic camera acquisition, standardization, and cleaning

IV. Explanation of cleaning and interpretation of COVID-19 data

Supplementary Tables

Table S1. Calendar dates for phases by year

Table S2. Population densities and geographic coverage for Centre County

Table S3. Traffic camera locations and descriptions

Table S4. Snowfall by year and phase

Supplementary Figures

Figure S1. Population density across Centre County

Figure S2. Daily mean radiance including full moon days

Figure S3. Splines fit to hourly traffic volume by individual cameras

Figure S4. Observed and predicted of hourly count data per camera through time

Supplementary Methods

I. Detailed timeline of events

Figure 1 in the main text details many of the key disease and policy events that are relevant to changes in policies and movement at the Penn State University Park campus and surrounding Centre County. However, here we provide citations for information in Figure 1 and additional events to give additional context to the policies and communication throughout.

Timeline of important disease and policy events local study periods:

Dec 31, 2019: China reports to WHO new pneumonia

Jan 20, 2020: 1st case in US[1]

[BASELINE PERIOD: Feb 18 – March 6]

February 12: first death in US (Santa Clara, CA, retrospectively identified) [2]

March 6: first 2 reported cases of COVID-19 in PA

[LOCAL POPULATION DECLINE: Feb 1 – March 6]

March 9: PSU spring break begins[3]

March 11: WHO declares COVID-19 a pandemic[4]; PSU announces no resident instruction following spring break[5]

March 16: PSU begins remote instruction[5]

March 19: all non-life-sustaining businesses ordered to close statewide by PA Governor[6]

March 21: enforcement of closure of non-life-sustaining businesses statewide[6]

[RED PHASE: Feb 1 – March 6]

March 28: Centre County receives stay at home order from PA Governor[7] ("Governor's Office: Official Website," n.d.)

April 3: PA Governor calls for "universal masking" [7]

April 9: PA schools officially closed through end of academic year

May 1: announcement from PA Governor that on May 8 PA will lift some restrictions in 24 counties in the Northcentral and Northwest health districts; all counties will remain in "red" until May 8, these counties will move to "yellow" on May 8 [7]

[YELLOW PHASE: Feb 1 – March 6]

May 8: PA lifts some restrictions in 24 counties in the Northcentral and Northwest health districts as determined by PA Governor in moving from "red" (Stay-at-home) to "yellow" (aggressive mitigation) [7]

May 28: US death toll from COVID-19 passes 100,000

[GREEN PHASE: May 29 – July 3]

May 29: Some of Northcentral and Northwest health districts (Fig S2) move from "yellow" to "green"

June 30: First PSU student death from COVID-19

July 1: Order of face coverings in public places in Pennsylvania[8]

II. Black Marble Pre-Processing Details

NASA VNP46A1 Daily At-Sensor Top of Atmosphere (TOA) Nighttime Radiance products were downloaded as Hierarchical Data Format 5 (HDF5) with a bounding box encompassing Centre County. Images are stored for days with timestamps from 00:00:00 and 23:59:59 for a given date, we make the assumption that the images are from the PM and thus reflect changes that happen during the day. However, it is possible the image is captured in the morning nighttime hours (00:00:00 to 07:00:00) and therefore reflect changes in policy from a day prior. CSV files for the radiance values, cloud mask values, and pixel latitude and longitude values were separated into separate Pandas DataFrames for each day.[9] We added missing dates (where there was no HDF5 data on a given day) and populated all data for dates added with Not a Number (NaN) values. We derived the plotting extent, transform, and array shape for the study area from the pixel latitude and longitude values DataFrame, for use with radiance data intermediate storage and export. We stored each day of radiance data as a NumPy array[10] in a Python dictionary, indexed by YEAR, MONTH, and DAY (e.g. radiance_data[YEAR][MONTH][DAY]). During the storage process, we masked the data for clouds based on the cloud mask values, scaled the data based on the scale factor (0.1), and shaped the NumPy array based on the derived spatial information from the previous pre-processing step. All pre-processing operations with radiance data were completed using Python 3.7.6.[11]

III. Details of traffic camera acquisition, standardization, and cleaning

Although we tried to capture images every 20 seconds, cameras could be offline or images weren't captured for each time interval. To standardize variation in images between cameras for a given hour,

the sum of vehicles (labeled over all images captured by that camera) from that hour is divided by the number of images captured and scaled by the number of images we expect if we were capturing one image per minute. This helps us to avoid overcounting cars stuck at traffic lights but likely underestimates total vehicle volume on interstates. Therefore, count data should be interpreted as relative increases rather than absolute.

Occasionally, live stream images from traffic cameras would freeze or cameras would go offline. The standardized hourly counts were cleaned to remove strings of zeros or integers (usually parked cars or frozen images). The incorrect data was replaced with NAs and predicted with the best fit generalized additive model (see Methods in Main Text; Fig S4).

III. Explanation of cleaning and interpretation of COVID-19 data

COVID-Cases are reported as daily cumulative cases. However, these cumulative cases did not monotonically increase over the observation period. When false positives are discovered, they are subtracted from the cumulative case total, resulting in some dates with negative new daily cases. To incorporate this data, we assumed that false positives were from the most recent previous day of reporting and no new cases occurred on the current day. There was not a way to determine the exact date of the false negative test, so this was our best estimate.

Centre County has one prison and six long term care homes with a bed capacity of ~500. Both of these settings are likely hotspots for COVID-19 transmission.[12] However, during the course of the study there were not major outbreaks at these locations and therefore the majority of Centre Count
In Huntingdon County, there was a prison outbreak and two days of data (with cases in excess of 30) were presumed to represent reporting from this outbreak. We removed these dates from the time series in Figure 1 in the main text to focus on community-driven transmission.

Supplementary Tables

Table S1. Calendar dates for phases by year. The timing of Spring Break and Commencement at Penn State effect the resident population of State College and surrounding Centre County. The timing of these events were matched across the years so that events in the pandemic year would be accurately compared to the same events/timing in previous year. All phases except for Population Decline and Local Restrictions have the same number of days in all years. For the Population Decline period, in all years but 2020, these dates included the Friday before spring break through the Sunday before classes resume. In 2020, this started on the Friday before spring break and extended to March 17 (Wednesday) while students were off campus but before non-essential businesses were closed. Dates are from the Office of the Registrar at Penn State.[3]

Year	**Baseline**	**Population Decline** (Spring Break)	**Local Restrictions**	**Red**	**Yellow** (corresponds to commencement)	**Green**
2016	Feb 12	March 4	March 14	March 25	May 6	May 27
2017	Feb 10	March 3	March 13	March 24	May 5	May 26
2018	Feb 9	March 2	March 12	March 23	May 4	May 25
2019	Feb 8	March 1	March 11	March 22	May 3	May 24
2020	Feb 14	March 6	March 18	March 27	May 8	May 29

Table S2. Population densities and geographic coverage for Centre County

Density category	**Pop range**	**Num pixels**	**Proportion of study area**	**Proportion of total population**
Excluded	(0,5]	-	-	0.014652247
Low [1]	(5,100]	7240	0.8618	0.218933101
Low-Mod [2]	(100,400]	823	0.096	0.218385905
Moderate [3]	(400,1000]	220	0.0279	0.186681017
Mod-High [4]	(1000,2000]	65	0.00986	0.151680971
High [4]	(2000,Inf]	39	0.0069	0.209666759

Table S3. Traffic camera locations and descriptions

ID	Latitude	Longitude	Ownership	Number of lanes	Road Type	Description
CAM02001CCTV2	40.810961	-78.075259	PennDOT (Pennsylvania Department of Transportation)	5	connector	Port Matilda US-322 By-Pass Westbound
CAM02002CCTV3	40.817144	-77.939841	PennDOT	7	connector	Grays Woods
CAM02003CCTV4	40.828613	-77.840339	PennDOT	6	connector	I-99/US-322 Interchange
CAM02005CCTV9	40.955477	-77.773749	PennDOT	3	connector	Milesburg Interchange West I-80 Exit 158 Eastbound / Alt US-220
CAM02006CCTV10	40.956061	-77.766277	PennDOT	3	connector	Milesburg Interchange East I-80 Exit 158 Westbound / PA-150
CAM02007CCTV13	40.795773	-77.820937	PennDOT	4	connector	US-322 E/O PA-26
CAM02009CCTV7	40.944571	-77.720918	PennDOT	2	connector	I-80 Eastbound Exit 161 Bellefonte Interchange (I-80 & I-99)
CAM02010CCTV11	40.80388	-78.063664	PennDOT	6	connector	Port Matilda I-99 Exit 61 Median
CAM02020CCTV24	40.829048	-77.804817	PennDOT	6	internal	Benner Pike
CAM02028CCTV32	40.812073	-77.9225	PennDOT	6	connector	Atherton and Valley Vista
CAM02033CCTV38	41.022703	-77.933981	PennDOT	4	connector	Snowshoe I-80 Exit 147
CAM02037CCTV43	40.807936	-77.895204	PennDOT	5	internal	Atherton and Vairo
CAM02038CCTV44	40.805031	-77.886886	PennDOT	5	internal	Atherton and North Hills
CAM02039CCTV45	40.796522	-77.872549	PennDOT	5	internal	Atherton and Park
CAM02040CCTV46	40.791787	-77.864971	PennDOT	4	internal	Atherton and W.College
CAM02042CCTV49	40.790894	-77.863812	PennDOT	4	internal	Atherton and W.Beaver
CAM02046CCTV52	40.817953	-77.900271	PennDOT	4	connector	I-99 at Exit 71 (Toftrees)
collegeTwp	40.811564	-77.830719	State College Township	4	connector	College Township Traffic camera
parkArboretum	40.804483	-77.864219	Penn State University	3	internal	Penn State University Park Ave.

Table S4. Snowfall by year and phase. Total snowfall in inches over each period and year; not snowfall accumulation.[13]

Year	Total snowfall		
	Baseline	**Population Decline**	**Local Restrictions**
2016	2.4”	0.1”	0.1”
2017	0.6”	5.3”	11.0”
2018	3.7”	1.4”	4.9”
2019	11.1”	4.9”	0.0”
2020	0.7”	0.0”	0.4”

Supplementary Figures

Figure S1. Population density across Centre County (per square kilometer). White areas are very low population densities (<1 person/ km2 or <5/km2). Radiance values for these pixels are minimally informative pixels and pixels with <5 were are excluded from the analysis. Population estimates are based on 1km rasters from WorldPop.[14,15] The boundaries of boroughs (townships) and county are shown.

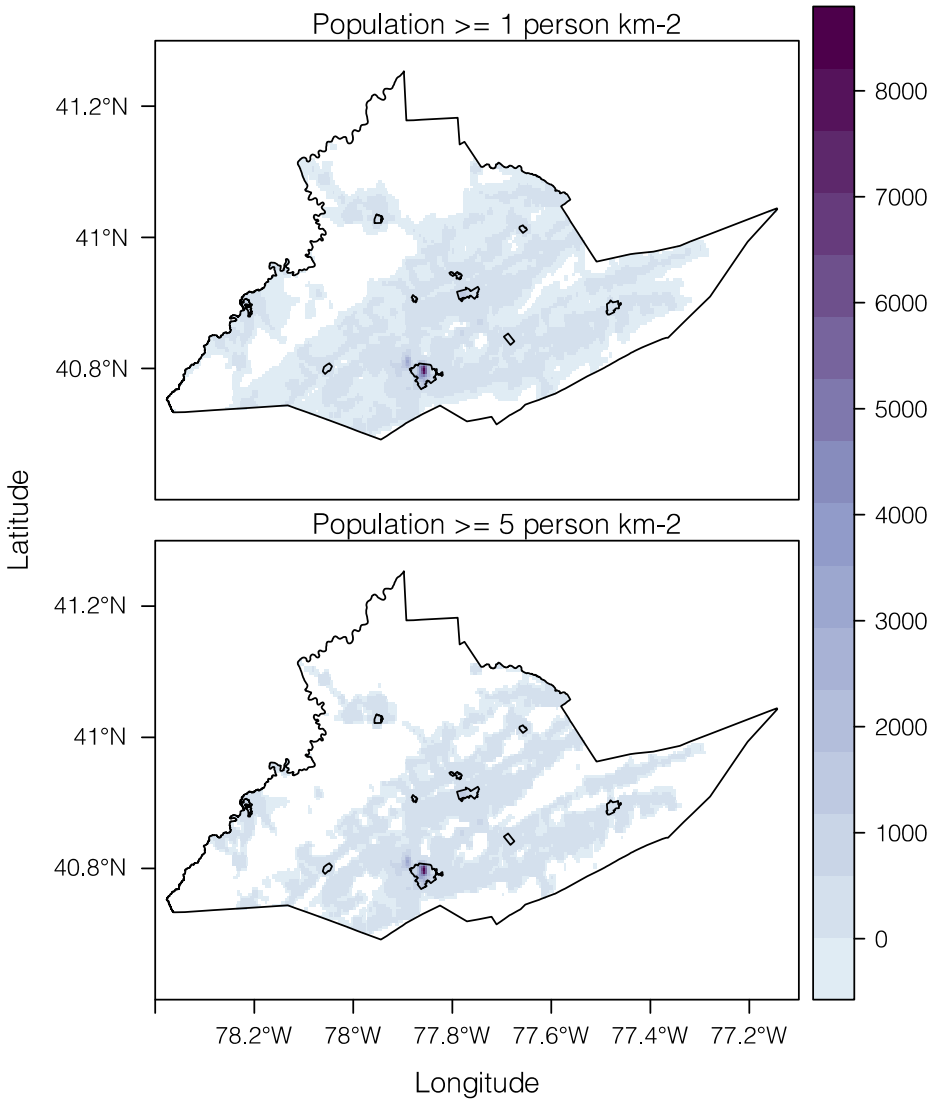


Figure S2. Daily mean radiance including full moon days. Mean and standard deviation of daily radiance values for all pixels (excluding very low density (<5 people km-2) areas). Points are colored in their respective phases for each year and full moon days (day of full moon and one day prior and three days following) are shown as open circles. Some dates had no cloud free pixels and there are no points for those dates. All other averages are the mean of all cloud free pixels and are not weighted for population.

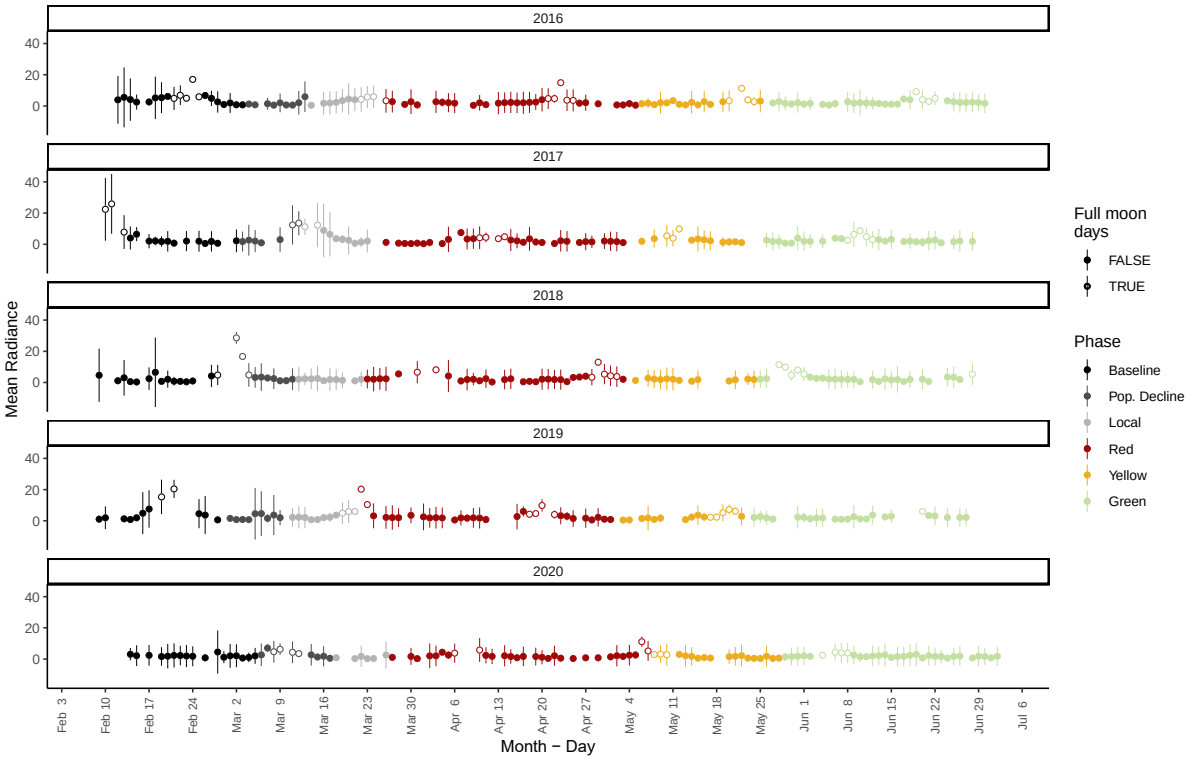


Figure S3. Splines fit to hourly traffic volume by individual cameras. For the axes: y- axis are the estimates of the best-fit GAM and x-axis is the hour of the day (range 0-23).

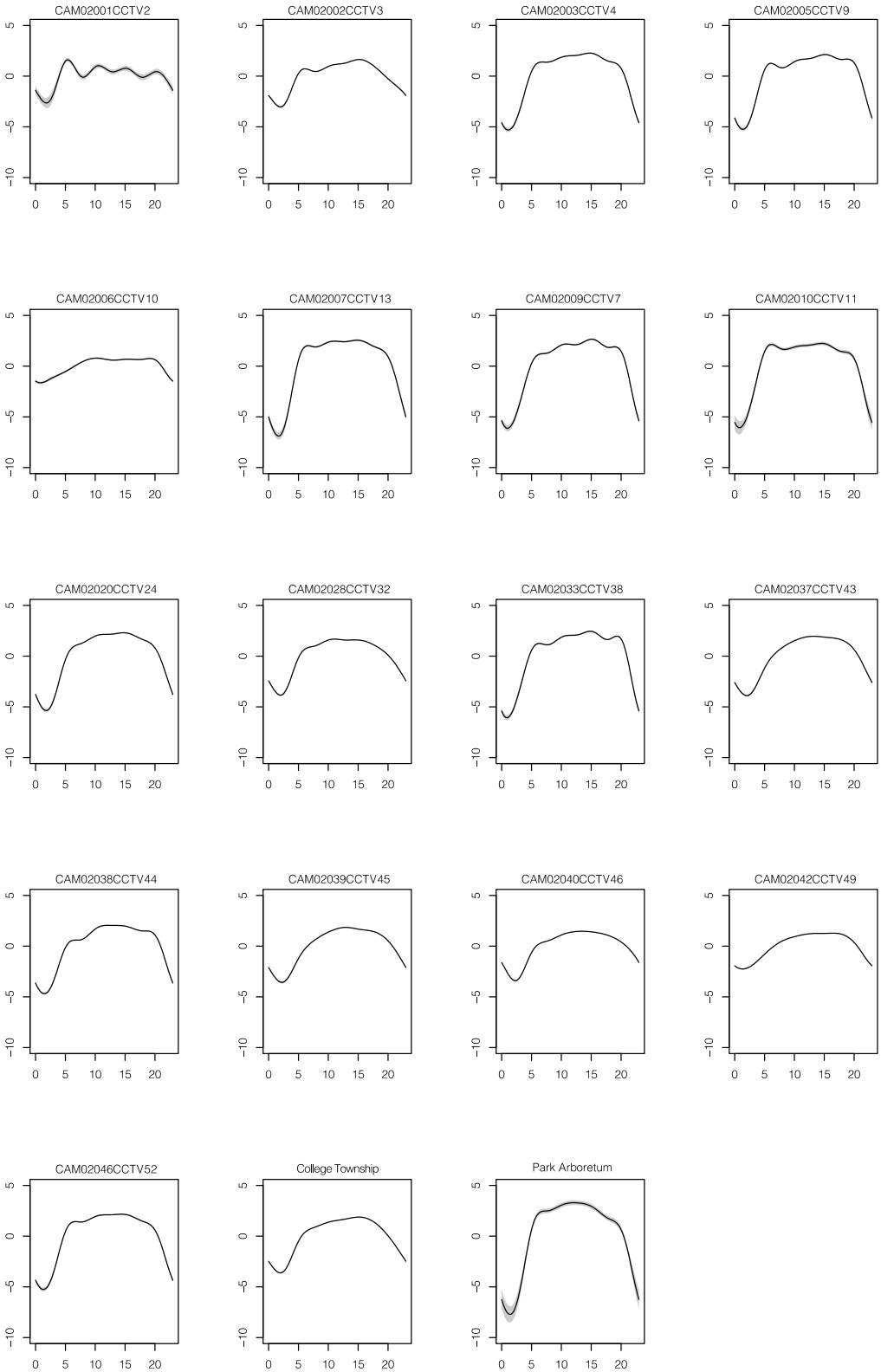
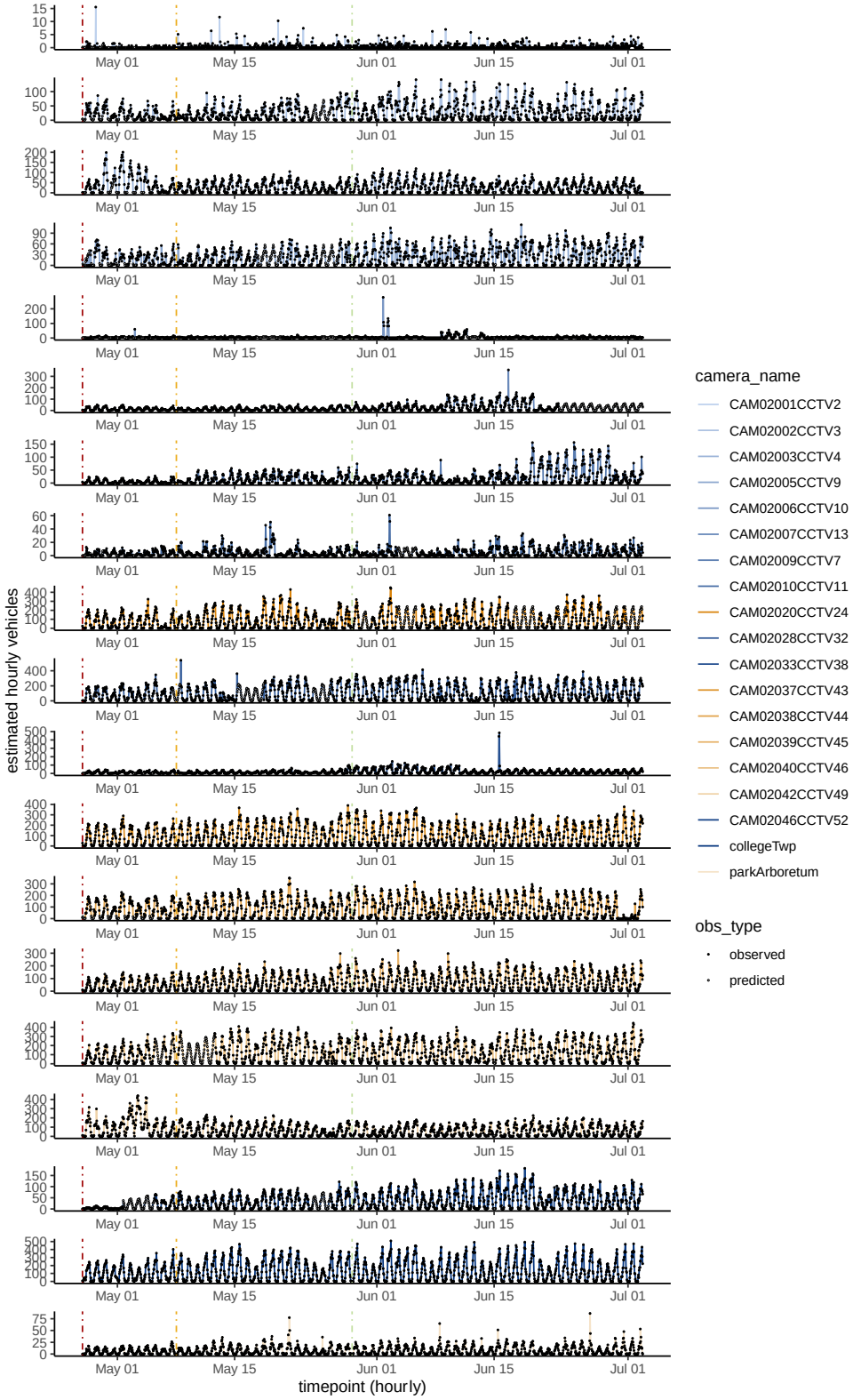


Figure S4. Observed and predicted hourly count data per camera through time. Predicted counts from the best-fit GAM are used to fill in gaps in observed data and are shown in open circles. Phases are indicated with vertical colored lines for Red, Yellow and Green Phases. Colors for each camera are consistent with Fig 3 in the main text – with shades of blue representing connector roads and orange shades representing internal roads. The range of the y-axis varies by camera.



Supplementary References

1. Holshue ML, DeBolt C, Lindquist S, et al. First case of 2019 novel coronavirus in the United States. *New England Journal of Medicine* 2020.
2. Fuller T, Baker M. Coronavirus death in California came weeks before first known US death. *NY Times https://www nytimes com/2020/04/22/us/coronavirus-first-united-states-death html Accessed* 2020; **6**.
3. PennState University Registrar. Academic Calendars. 2020. https://registrar.psu.edu/academic-calendars/ (accessed 08/01/2020.
4. World Health O. WHO Director-General's opening remarks at the media briefing on COVID-19, 2020.
5. Barron EJ. All Penn State classes to take place remotely beginning March 16. Penn State University. 2020.
6. Governor T Wolf. Order of the Govenor of the Commonwealth of Pennsylvania regarding the closure of all businesses that are not life sustaining. In: Governor Oot, editor.: Commonwealth of Pennsylvania; 2020.
7. Wolf GT. Process to Reopen Pennsylvania. 2020. https://www.governor.pa.gov/process-to-reopen-pennsylvania/.
8. Levine R. Order of the Secretary of the Pennsylvania Department of Health Requiring Universal Face Coverings. In: Health PDo, editor.; 2020.
9. McKinney W. Data structures for statistical computing in python. Proceedings of the 9th Python in Science Conference; 2010: Austin, TX; 2010. p. 51-6.
10. Oliphant TE. 2006: Trelgol Publishing, A guide to NumPy.
11. Python Software F. Python Language Reference, version 3.7.6. 2019.
12. Department of H. COVID-19 Long-Term Care Facilities Data for Pennsylvania. 2020.
13. Penn State Joel N. Myers Weather Center. State College Daily Weather Summary and Statistics. 2020. http://www.meteo.psu.edu/~wjs1/wxstn/wxstn.htm (accessed 2020/08/01.
14. WorldPop (www.worldpop.org - School of Geography and Environmental Science UoSDoGaG, University of Louisville; Departement de Geographie, Universite de Namur) and Center for International Earth Science Information Network (CIESIN), Columbia University. Global High Resolution Population Denominators Project - Funded by The Bill and Melinda Gates Foundation (OPP1134076). In: www.worldpop.org, editor.; 2018.
15. Tatem AJ. WorldPop, open data for spatial demography. *Scientific data* 2017; **4**.