

Investigating the effects of absolute humidity and human encounters on transmission of COVID-19 in the United States

Gary Lin¹, Alisa Hamilton¹, Oliver Gatalo¹, Fardad Haghpanah¹, Takeru Igusa^{2,3,4}, Eili Klein^{1,5,6},
For the CDC MInD-Healthcare Network

1. Center for Disease Dynamics, Economics & Policy, Silver Spring, MD, USA
2. Department of Civil and Systems Engineering, Johns Hopkins University, Baltimore, MD, USA
3. Department of Earth and Planetary Sciences, Johns Hopkins University, Baltimore, MD, USA
4. Center for Systems Science and Engineering, Johns Hopkins University, Baltimore, MD, USA
5. Department of Emergency Medicine, Johns Hopkins University, Baltimore, MD, USA
6. Department of Epidemiology, Johns Hopkins University, Baltimore, MD, USA

Corresponding author: Eili Klein

Keywords: COVID-19, SARS-Cov-2, Humidity, Non-essential visits

Abstract

Background

Mounting evidence suggests that the primary mode of transmission of SARS-CoV-2 is aerosolized transmission from close contact with infected individuals. Even though transmission is a direct result of human encounters, environmental conditions, such as lower humidity, may enhance aerosolized transmission risks similar to other respiratory viruses such as influenza.

Methods

We utilized dynamic time warping to cluster all 3,137 counties in the United States based on temporal data on absolute humidity from March 10 to September 29, 2020. We then used a multivariate generalized additive model (GAM) combining data on human mobility derived from mobile phone data with humidity data to identify the potential effect of absolute humidity and mobility on new daily cases of COVID-19 while considering the temporal differences between seasons.

Results

The clustering analysis found ten groups of counties with similar humidity levels. We found a significant negative effect between increasing humidity and new cases of COVID-19 in most regions, particularly in the period from March to July. The effect was greater in regions with generally lower humidity in the Western, Midwest, and Northeast regions of the US. In the two regions with the largest effect, a 1 g/m³ increase of absolute humidity resulted in a 0.21 and 0.15 decrease in cases. The effect of mobility on cases was positive and significant across all regions in the July-Sept time period, though the relationship in some regions was more mixed in the March to June period.

Conclusions

We found that increasing humidity played an important role in falling cases in the spring, while increasing mobility in the summer contributed more significantly to increases in the summer. Our findings suggest that, similar to other respiratory viruses, the decreasing humidity in the winter is likely to lead to an increase in COVID-19 cases. Furthermore, the fact that mobility data were positively correlated suggests that efforts to counteract the rise in cases due to falling humidity can be effective in limiting the burden of the pandemic.

Introduction

To date, the coronavirus disease 2019 (COVID-19) pandemic has claimed over 225,000 lives in the United States alone, with more than 8.6 million confirmed cases (1). The primary mode of transmission of the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) that causes COVID-19 is close contact with infected individuals (2, 3). Growing evidence suggests aerosols (4, 5), defined as particulates less than 5 μm in diameter (6, 7), likely play the most important role in transmission (8). The first surge of cases in the United States was seen in the late winter of 2020, but the vast majority of the epidemic to date has occurred during the spring and summer. As the fall season begins in most areas of the northern hemisphere, the weather will trend towards colder and drier. Falling levels of absolute humidity have been shown to increase transmission rates of other respiratory viruses, such as influenza (9), posing significant concern regarding potential increases in the number of COVID-19 cases in the fall and winter.

While several studies have suggested there may be a relationship between climatic factors such as temperature and/or humidity and COVID-19 (10–16), uncertainty still remains in the precise environmental and biological mechanisms (e.g., relative humidity, vapor pressure, and temperature) behind aerosol and droplet transmissions and viral survival of SARS-CoV-2 (17). In influenza, lower atmospheric moisture has been shown to increase the production of aerosol nuclei and viral survival time (9), which translates to higher risks of airborne transmission. Other climatic factors that may impact transmission include temperature and air quality (18, 19); however, absolute humidity can still provide a surrogate measure for indoor air moisture and temperature (20).

Initial efforts to slow the spread of COVID-19 focused on reducing contacts between individuals through social-distancing measures such as large-scale lockdowns, which were significantly associated with reductions in cases (21). However, as the initial lockdowns were lifted and movement of individuals increased, the correlation between mobility and case growth rates weakened (22). Though some counties and states saw increases in cases, others saw decreases without corresponding increases in movement by any metric. Thus, other factors besides mobility patterns are likely to be drivers of transmission.

Analyses of the factors influencing COVID-19 have used either climate data (19, 23–25) or human mobility data (21), but no study that we know of has considered changes in both climate and human mobility on COVID-19 outbreaks in the United States. Preliminary studies have investigated these effects in China but did not consider the varying sensitivities to humidity for different climatological regimes, which can lead to weak detection in humidity impacts on transmission (26). Understanding the potential for climatic factors to increase transmission in the fall and winter is crucial for developing policies to combat the spread of the SARS-CoV-2. While the interaction between environmental factors and human encounters is complex, accounting for this relationship is necessary for deciding business and school reopenings. Furthermore, as indoor gatherings, which typically increase in frequency and size in the winter, are one of the largest risk factors for transmission (7, 27), greater understanding is needed as to the added risk of changes in weather to aid decisions on when to restrict gatherings or implement mandates for protective face coverings. In this study, we assess the relative impact of absolute humidity in different climatological regimes and human mobility on reported cases of COVID-19 across counties in the US.

Methods

Data sources

Daily average absolute humidity for each US county, excluding territories, was calculated using temperature and dewpoint data from the National Centers for Environmental Information (28) at the National Oceanic and Atmospheric Administration (NOAA). Time series data for the year 2020 from US weather stations were acquired from the NOAA Global Summary of the Day Index (29). Weather stations were mapped using latitude and longitude to corresponding counties using the Federal Communications Commission (FCC) Census Block API (30). For counties without a weather station, we used data from the nearest station, which was calculated based on distance from the county's spatial centroid using the haversine formula. For counties with multiple stations, data were averaged across all stations in a county. Absolute humidity was calculated using average daily temperature and average daily dew point (see (31)).

Data on visits to non-essential businesses from March 10, 2020 to September 29, 2020 was obtained from the Unacast Social Distancing Scorecard (32). We specifically utilized the metric that measures visits to non-essential places by comparing the rate on the county-level to the national average. This temporal measure allowed us to compare across counties. The data on non-essential visits was calculated as the percent difference from before policy interventions (e.g., shelter-in-place orders) began to impact movement.

Confirmed case data were extracted from the Johns Hopkins Center for Systems Science and Engineering (1), and the population size data for each county were obtained from the US Census Bureau (33) for 3,137 counties from March 10, 2020 to September 29, 2020. Daily cases were obtained from the confirmed case count by taking a simple difference between the days. Any data incongruencies, such as negative case counts, were omitted in our analysis.

Statistical Analysis

The United States is geographically large, and the timing and magnitude of changes in absolute humidity can vary widely across regions. In order to account for regional differences in humidity, we utilized a partitioned clustering algorithm with dynamic time warping (DTW) similarity measurements (34) to classify the absolute humidity profile for all observed counties into ten exclusive clusters. Clustering allowed us to designate groups of counties based on temporal, climatological regimes and to stratify different absolute humidity patterns, which reduces group-level effects and enhances the independence of the data points. The DTW clustering of absolute humidity was conducted on a larger set of 3,137 counties. In the regression analysis, we only examined counties that had more than twenty cumulative confirmed cases by September 29th, 2020, or a population of more than 50,000 people. We excluded any days that had fewer than 20 cumulative confirmed cases within each county because early transmission dynamics had a high rate of undetected cases (35), making the data unreliable for this analysis. The final dataset used in the regression analysis included 987 counties. We assessed the results of the model over three time periods in 2020: (1) the entire duration of the dataset (March 10 to Sept 29), (2) spring when humidity increases (March 10 to June 30), and (3) the summer months when humidity is relatively steady (July 1 to Sept 29).

For each humidity cluster, we conducted three multivariate regressions using a generalized additive model (GAM) that assessed the time-weighted association between absolute humidity and non-essential visits with the number of new coronavirus cases (Eqns. 1-3). GAMs are semiparametric models that have been used extensively in assessing impacts of climatic variables

and health outcomes (36, 37) and allowed us to estimate a functional variability with basis splines (38). The model we used is described in the following equation,

$$\log(Y_{it}) = \log(N) + \alpha + f(t) + \beta_1 AH_{i(t-\delta)} + \beta_2 NV_{i(t-\delta)} + \gamma_i + \epsilon_{it} \quad (1)$$

where Y_{it} , is the number of daily COVID-19 cases for county i at time t , $\log(N)$ is an offset term to control for population-size, and α is the intercept. Absolute humidity, $AH_{i(t-\delta)}$, and non-essential visitations, $NV_{i(t-\delta)}$, were smoothed using a 7-day moving average and lagged by δ days. For our study, we assumed that δ was equal to 14 days, which is based on previous studies investigating lagged effects due to the incubation period of COVID-19 (39). Fixed effects, defined as γ_i , are included for each county in our analysis to control for variations between counties. The operator $f(\cdot)$ is a spline basis function that is used to fit nonlinearities attributed to unobserved time-varying effects. Since we were predicting daily cases, we assumed the response variable is a Poisson distributed random variable with a log-transformed link function. Standard errors were calculated for the estimated linear coefficients β_1 and β_2 .

To test for robustness, we conducted additional regressions on the absolute humidity and non-essential visits predictors individually with fixed effects for counties. Specifically, for each humidity cluster, we fitted a GAM with non-essential visitations and absolute humidity as linear predictors for new daily cases, as described in Equations (2) and (3).

$$\log(Y_{it}) = \log(N) + \alpha + f(t) + \beta_1 NV_{i(t-\delta)} + \gamma_i + \epsilon_{it} \quad (2)$$

$$\log(Y_{it}) = \log(N) + \alpha + f(t) + \beta_1 AH_{i(t-\delta)} + \gamma_i + \epsilon_{it} \quad (3)$$

The multiple regression analyses were compared to demonstrate robustness in the coefficient estimates and the relative effects of absolute humidity and non-essential visitations. The analysis using GAM was conducted using the *mgcv* packages in R (Version 4.0.2).

Results

Using the dynamic time warping (DTW) algorithm to characterize similarities of humidity profiles, we were able to partition all counties in our humidity dataset into ten exclusive groups. The clustering of humidity coincided with various geographic regions in the US (Figure 1A), which supports the notion that climate is spatially dependent. The cluster with the lowest average absolute humidity is Group 7 (398 counties), which is primarily located in the western region of the US. Group 6 (230 counties) had the highest average absolute humidity and is located on the southern coast bordering the Gulf of Mexico (see Figure 1B). Group 4 had the most varying geography, with counties from the West Coast, Midwest, and North East United States clustered together using the DTW algorithm.

The absolute humidity in counties that belonged to Groups 4 and 7 consistently had the largest negative effect on new cases for all three time periods (Tables 1-3). For the 7-month observation, Group 4 showed that a 1 g/m^3 increase of absolute humidity resulted in a 0.21 decrease of new cases per day, while Group 7 had a 0.15 decrease in cases. When compared to the observations between March 10 and June 30, Group 4 still had a reduction of 0.22 new cases for each unit increase in absolute humidity, and Group 7 had a 0.18 decrease in cases. After July 1, absolute

humidity had a relatively smaller impact on both groups, with Group 7 decreasing by 0.15 cases and Group 4 decreasing by 0.09 cases per day for each unit increase in absolute humidity. The cluster with the highest relative contribution to daily cases due to non-essential visitations for the entire six-month observation was Group 2 (Table 1). Group 9 had the largest increase in new cases due to an increase in non-essential visitations before July (Table 2), while Group 5 had the largest increase after July (Table 3). Generally, we see a stronger, positive association with non-essential visitations for all groups, except Group 9, for July to September (Table 3) than March to June (Table 2).

The coefficient estimates and fitted plots for the additional GAM to show robustness are shown in the supplementary appendix (Tables S1-S10). For all models, the estimated degrees of freedom for the time-varying basis spline were approximately 8 (see Figures S2 in supplementary appendix for fits). The coefficient estimates for Groups 4 and 7 were consistently the largest for those additional fits, indicating robustness in the results.

Discussion

The approaching winter in the United States has been of great concern because of the link between wintertime and increasing transmission of respiratory viruses such as influenza and the other pandemic coronaviruses, MERS-CoV and SARS-CoV-1 (9, 40–43). SARS-CoV-2 is a novel human virus, and there remains great uncertainty as to the extent that disease transmission will increase over the next several months. Here we found that the relative effect of absolute humidity on transmissions has so far been significant and was greatest in the Western, Midwest, and Northeast regions of the United States, which were clustered into the driest climatological regimes. This study outcome supports the notion that absolute humidity affects transmission risk of respiratory viruses in regions that are arid and dry (44). However, this effect was less noticeable for more humid regions, such as the coastal counties and the southern regions of the US (Figure 2). We also found that non-essential visitations had a stronger, more positive impact on emergence of cases during the summer months. This suggests that an increase in contacts due to human mobility had a relatively more extensive role in case growth when absolute humidity is steady and higher.

As the absolute humidity decreases during winter seasons, this suggests that COVID-19 may increase significantly over the next couple months. Increasing COVID-19 cases will pose an issue for many healthcare systems, which in normal years are typically stretched thin from regular seasonal infections, such as influenza. Furthermore, seasonal changes in human behavior may also impact the number of new cases and hospitalizations since people are more likely to occupy indoor spaces for longer durations when outdoor temperatures decrease, thus increasing the risk of transmission. With lower absolute humidity levels, relative humidity indoors will also be lower, causing higher susceptibility to airborne diseases (45).

The mechanisms through which changes in absolute humidity might affect transmission of respiratory viruses are not well understood. Several studies have shown that as absolute humidity falls, survival times for enveloped viruses increases, including other coronaviruses (9, 20, 46), though there are some suggestions that these effects may be nonlinear (47). Our findings supported this nonlinear relationship since the log-linear effects between humidity and case growth varied between climatological regimes. One potential factor may be that increasing humidity results in degradation of the capsid envelope leading to viral degradation (46, 48). Another possibility may be that higher absolute humidity reduces the virus's binding capacity,

reducing the potential infectivity (48). Further studies are needed to refine the mechanisms by which humidity changes may alter transmission of SARS-CoV-2 and the related functional form. However, the results of these prior studies and the association found in the current study suggest that increased humidification of indoor air in high transmission settings may help decrease the impact of COVID-19.

While our findings suggest that declines in absolute humidity are likely to contribute to increased transmission of SARS-CoV-2, our findings also demonstrate the importance of contact and movement of individuals. For example, in the 2009 influenza pandemic, the increased contact patterns that occur in the fall likely combined with falling humidity to drive transmission, which resulted in the peak of infections occurring significantly earlier than other years. Given the uncertainty of the nonlinear effects of humidity on transmission, continued restrictions on large gatherings and mandates for face-coverings are likely crucial in the coming months to reducing the toll of the pandemic. In addition, the uncertainty regarding the role of children in transmission (49–51) suggests that caution in opening schools is warranted as the potential for transmission increases, which contributes to seasonality effects. While studies linking schools to outbreaks to date have been limited, few have occurred during the winter when transmission may be higher. Studies assessing infections of children are urgently needed in areas where schools have opened.

Studies of prior viruses and preliminary studies of the SARS-CoV-2 virus underpin the theoretical connection between humidity and transmission; however, limitations remain in assessing this association. One limitation of this study includes the spatial coupling of populations due to travel and commute, which may alter the transmission patterns within counties. Spatial clustering of the analysis by trends in humidity should reduce this effect somewhat. A second limitation is an intra-county heterogeneity of the population (i.e., income, demographics, and social status), which we attempted to control for with fixed effects in our model. Third, as with most COVID-19 analyses on retrospective data, heterogeneity in testing at the county level likely results in differences in the correlation between reported results and actual cases. Fourth, other exogenous events, such as evacuation-related activities related to natural disasters or mass-gatherings in several cities this summer (52), may bias the connection between humidity and transmission. Finally, there may also be time-varying trends in other factors that affect case detection rates, though GAMs reduce the potential impact of these factors as they account for time-varying trends related to unobservable factors external to the independent variables in the model.

The upcoming winter in the United States and other temperate regions in the northern hemisphere are likely to increase transmission of SARS-CoV-2 due to falling humidity. In the US, these impacts are likely to be largest in the Western, Midwest, and Northeast regions of the United States that have lower humidity trends relative to the rest of the country. Given that much of our understanding of transmission has occurred during periods of higher humidity, significant uncertainty remains as to the impact of this increased transmission on morbidity and mortality of COVID-19. However, underestimating the potential for increased transmission could lead to needless deaths, particularly in nursing homes and long-term care facilities, where the possibility exists that infection control procedures in the summer may not be adequate in the winter and limits on visitors and regular screening of staff should be instituted sooner rather than later.

Acknowledgments

This work was funded by the Centers for Disease Control and Prevention (CDC) MInD-Healthcare Program (Grant Numbers U01CK000589, 1U01CK000536, and contract number 75D30120P07912). The funders had no role in the design, data collection and analysis, decision to publish, or preparation of the manuscript.

References

1. E. Dong, H. Du, L. Gardner, An interactive web-based dashboard to track COVID-19 in real time. *The Lancet Infectious Diseases*. **20**, 533–534 (2020).
2. J. Liu, X. Liao, S. Qian, J. Yuan, F. Wang, Y. Liu, Z. Wang, F.-S. Wang, L. Liu, Z. Zhang, Community Transmission of Severe Acute Respiratory Syndrome Coronavirus 2, Shenzhen, China, 2020. *Emerging Infectious Diseases*. **26**, 1320–1323 (2020).
3. J. F.-W. Chan, S. Yuan, K.-H. Kok, K. K.-W. To, H. Chu, J. Yang, F. Xing, J. Liu, C. C.-Y. Yip, R. W.-S. Poon, H.-W. Tsoi, S. K.-F. Lo, K.-H. Chan, V. K.-M. Poon, W.-M. Chan, J. D. Ip, J.-P. Cai, V. C.-C. Cheng, H. Chen, C. K.-M. Hui, K.-Y. Yuen, A familial cluster of pneumonia associated with the 2019 novel coronavirus indicating person-to-person transmission: a study of a family cluster. *The Lancet*. **395**, 514–523 (2020).
4. CDC, Coronavirus Disease 2019 (COVID-19). *Centers for Disease Control and Prevention* (2020), (available at <https://www.cdc.gov/coronavirus/2019-ncov/more/scientific-brief-sars-cov-2.html>).
5. S. Tang, Y. Mao, R. M. Jones, Q. Tan, J. S. Ji, N. Li, J. Shen, Y. Lv, L. Pan, P. Ding, X. Wang, Y. Wang, C. R. MacIntyre, X. Shi, Aerosol transmission of SARS-CoV-2? Evidence, prevention and control. *Environ Int*. **144**, 106039 (2020).
6. J. Ma, X. Qi, H. Chen, X. Li, Z. Zhang, H. Wang, L. Sun, L. Zhang, J. Guo, L. Morawska, S. A. Grinshpun, P. Biswas, R. C. Flagan, M. Yao, Coronavirus Disease 2019 Patients in Earlier Stages Exhaled Millions of Severe Acute Respiratory Syndrome Coronavirus 2 Per Hour. *Clin Infect Dis*, doi:10.1093/cid/ciaa1283.
7. P. de Man, S. Paltansing, D. S. Y. Ong, N. Vaessen, G. van Nielen, J. G. M. Koeleman, Outbreak of Coronavirus Disease 2019 (COVID-19) in a Nursing Home Associated With Aerosol Transmission as a Result of Inadequate Ventilation. *Clin Infect Dis*, doi:10.1093/cid/ciaa1270.
8. H. S. Rahman, M. S. Aziz, R. H. Hussein, H. H. Othman, S. H. Salih Omer, E. S. Khalid, N. A. Abdulrahman, K. Amin, R. Abdullah, The transmission modes and sources of COVID-19: A systematic review. *International Journal of Surgery Open*. **26**, 125–136 (2020).
9. J. Shaman, M. Kohn, Absolute humidity modulates influenza survival, transmission, and seasonality. *Proceedings of the National Academy of Sciences*. **106**, 3243–3248 (2009).
10. Y. Wu, W. Jing, J. Liu, Q. Ma, J. Yuan, Y. Wang, M. Du, M. Liu, Effects of temperature and humidity on the daily new cases and new deaths of COVID-19 in 166 countries. *Science of The Total Environment*. **729**, 139051 (2020).
11. M. Kato, T. Sakihama, Y. Kinjo, D. Itokazu, Y. Tokuda, “Effect of Climate on COVID-19 Incidence: A Cross-Sectional Study in Japan” (SSRN Scholarly Paper ID 3612114, Social Science Research Network, Rochester, NY, 2020), , doi:10.2139/ssrn.3612114.

12. C. A. Christophi, M. S. Prieto, F.-Y. Lan, M. Delgado, V. Efthymiou, G. C. Gaviola, A. Hadjivasilis, Y.-T. Hsu, A. Kyprianou, I. Lidoriki, C.-F. Wei, F. Rodriguez-Artalejo, S. N. Kales, "Ambient Temperature and Subsequent COVID-19 Mortality in the OECD Countries and Individual United States" (SSRN Scholarly Paper ID 3612112, Social Science Research Network, Rochester, NY, 2020), , doi:10.2139/ssrn.3612112.
13. A. Meyer, R. Sadler, C. Faverjon, A. R. Cameron, M. Bannister-Tyrrell, Evidence That Higher Temperatures Are Associated With a Marginally Lower Incidence of COVID-19 Cases. *Front. Public Health*. **8** (2020), doi:10.3389/fpubh.2020.00367.
14. E. Steiger, T. Mussnug, L. E. Kroll, *medRxiv*, in press, doi:10.1101/2020.07.15.20154476.
15. D. Kifer, D. Bugada, J. Villar-Garcia, I. Gudelj, C. Menni, C. H. Sudre, F. Vuckovic, I. Ugrina, L. F. Lorini, S. Bettinelli, N. Ughi, A. Maloberti, O. Epis, C. Giannattasio, C. Rossetti, L. Kalogjera, J. Persec, L. Ollivere, B. Ollivere, H. Yan, T. Cai, G. Aithal, C. Steves, A. Kantele, M. Kajova, O. Vapalahti, A. Sajantila, R. Wojtowicz, W. Wierzbza, Z. Krol, A. Zaczynski, K. Zycinska, M. Postula, I. Luksic, R. Civljak, A. Markotic, C. Mahnkopf, A. Markl, J. Brachmann, B. Murray, S. Ourselin, J. Pascual, A. M. Valdes, M. Posso, J. Horcajada, X. Castells, M. Allegri, D. Primorac, T. Spector, C. Barrios, G. Lauc, *medRxiv*, in press, doi:10.1101/2020.07.11.20147157.
16. S. T. Sehra, J. D. Saliccioli, D. J. Wiebe, S. Fundin, J. F. Baker, Maximum Daily Temperature, Precipitation, Ultraviolet Light, and Rates of Transmission of Severe Acute Respiratory Syndrome Coronavirus 2 in the United States. *Clin Infect Dis*, doi:10.1093/cid/ciaa681.
17. S. Yuan, S.-C. Jiang, Z.-L. Li, Do Humidity and Temperature Impact the Spread of the Novel Coronavirus? *Front. Public Health*. **8** (2020), doi:10.3389/fpubh.2020.00240.
18. S. Lolli, Y.-C. Chen, S.-H. Wang, G. Vivone, Impact of meteorological conditions and air pollution on COVID-19 pandemic transmission in Italy. *Scientific Reports*. **10**, 16213 (2020).
19. M. M. Sajadi, P. Habibzadeh, A. Vintzileos, S. Shokouhi, F. Miralles-Wilhelm, A. Amoroso, Temperature, Humidity, and Latitude Analysis to Estimate Potential Spread and Seasonality of Coronavirus Disease 2019 (COVID-19). *JAMA Netw Open*. **3** (2020), doi:10.1001/jamanetworkopen.2020.11834.
20. L. C. Marr, J. W. Tang, J. Van Mullekom, S. S. Lakdawala, Mechanistic insights into the effect of humidity on airborne influenza virus survival, transmission and incidence. *J R Soc Interface*. **16** (2019), doi:10.1098/rsif.2018.0298.
21. H. S. Badr, H. Du, M. Marshall, E. Dong, M. M. Squire, L. M. Gardner, Association between mobility patterns and COVID-19 transmission in the USA: a mathematical modelling study. *The Lancet Infectious Diseases*. **0** (2020), doi:10.1016/S1473-3099(20)30553-3.

22. O. Gatalo, K. Tseng, A. Hamilton, G. Lin, E. Klein, Associations between phone mobility data and COVID-19 cases. *The Lancet Infectious Diseases*. **0** (2020), doi:10.1016/S1473-3099(20)30725-8.
23. J. Wang, K. Tang, K. Feng, X. Lin, W. Lv, K. Chen, F. Wang, High Temperature and High Humidity Reduce the Transmission of COVID-19. *SSRN* (2020), doi:10.2139/ssrn.3551767.
24. H. Qi, S. Xiao, R. Shi, M. P. Ward, Y. Chen, W. Tu, Q. Su, W. Wang, X. Wang, Z. Zhang, COVID-19 transmission in Mainland China is associated with temperature and humidity: A time-series analysis. *Science of The Total Environment*. **728**, 138778 (2020).
25. J. Xie, Y. Zhu, Association between ambient temperature and COVID-19 infection in 122 cities from China. *Science of The Total Environment*. **724**, 138201 (2020).
26. C. Poirier, W. Luo, M. S. Majumder, D. Liu, K. D. Mandl, T. A. Mooring, M. Santillana, The role of environmental factors on transmission rates of the COVID-19 outbreak: an initial assessment in two spatial scales. *Scientific Reports*. **10**, 17002 (2020).
27. R. K. Bhagat, M. S. D. Wykes, S. B. Dalziel, P. F. Linden, Effects of ventilation on the indoor spread of COVID-19. *Journal of Fluid Mechanics*. **903** (2020), doi:10.1017/jfm.2020.720.
28. National Centers for Environmental Information (NCEI), (available at <https://www.ncei.noaa.gov/>).
29. Global Surface Summary of the Day - GSOD - NOAA Data Catalog, (available at <https://data.noaa.gov/dataset/dataset/global-surface-summary-of-the-day-gsod>).
30. Census Block Conversions API. *Federal Communications Commission* (2011), (available at <https://www.fcc.gov/census-block-conversions-api>).
31. O. A. Alduchov, R. E. Eskridge, Improved Magnus Form Approximation of Saturation Vapor Pressure. *J. Appl. Meteor*. **35**, 601–609 (1996).
32. Covid-19 Social Distancing Scoreboard — Unacast, (available at <https://www.unacast.com/covid19/social-distancing-scoreboard>).
33. 2018 ACS 1-year Estimates. *The United States Census Bureau*, (available at <https://www.census.gov/programs-surveys/acs/technical-documentation/table-and-geography-changes/2018/1-year.html>).
34. D. J. Berndt, J. Clifford, Using Dynamic Time Warping to Find Patterns in Time Series, 12.
35. J. D. Silverman, N. Hupert, A. D. Washburne, Using influenza surveillance networks to estimate state-specific prevalence of SARS-CoV-2 in the United States. *Science Translational Medicine*. **12** (2020), doi:10.1126/scitranslmed.abc1126.

36. F. Dominici, A. McDermott, S. L. Zeger, J. M. Samet, On the Use of Generalized Additive Models in Time-Series Studies of Air Pollution and Health. *American Journal of Epidemiology*. **156**, 193–203 (2002).
37. K. Ravindra, P. Rattan, S. Mor, A. N. Aggarwal, Generalized additive models: Building evidence of air pollution, climate change and human health. *Environment International*. **132**, 104987 (2019).
38. T. J. Hastie, R. J. Tibshirani, Generalized additive models. *Statistical Science*. **1**, 297–310 (1986).
39. S. A. Lauer, K. H. Grantz, Q. Bi, F. K. Jones, Q. Zheng, H. R. Meredith, A. S. Azman, N. G. Reich, J. Lessler, The Incubation Period of Coronavirus Disease 2019 (COVID-19) From Publicly Reported Confirmed Cases: Estimation and Application. *Annals of Internal Medicine*. **172**, 577–582 (2020).
40. E. G. Gardner, D. Kelton, Z. Poljak, M. Van Kerkhove, S. von Dobschuetz, A. L. Greer, A case-crossover analysis of the impact of weather on primary cases of Middle East respiratory syndrome. *BMC Infect Dis*. **19**, 113 (2019).
41. K. H. Chan, J. S. M. Peiris, S. Y. Lam, L. L. M. Poon, K. Y. Yuen, W. H. Seto, The Effects of Temperature and Relative Humidity on the Viability of the SARS Coronavirus. *Advances in Virology*. **2011**, 1–7 (2011).
42. J. Yuan, H. Yun, W. Lan, W. Wang, S. G. Sullivan, S. Jia, A. H. Bittles, A climatologic investigation of the SARS-CoV outbreak in Beijing, China. *American Journal of Infection Control*. **34**, 234–236 (2006).
43. A. Altamimi, A. E. Ahmed, Climate factors and incidence of Middle East respiratory syndrome coronavirus. *Journal of Infection and Public Health*. **13**, 704–708 (2020).
44. A. A. R. V, V. R, F. Haghghat, The contribution of dry indoor built environment on the spread of Coronavirus: Data from various Indian states. *Sustain Cities Soc*. **62**, 102371 (2020).
45. G. Science, An Overview on the Role of Relative Humidity in Airborne Transmission of SARS-CoV-2 in Indoor Environments. *Aerosol Air Qual. Res*. **20**, 1856–1861 (2020).
46. J. McDevitt, S. Rudnick, M. First, J. Spengler, Role of absolute humidity in the inactivation of influenza viruses on stainless steel surfaces at elevated temperatures. *Appl Environ Microbiol*. **76**, 3943–3947 (2010).
47. J. Shaman, E. Goldstein, M. Lipsitch, Absolute Humidity and Pandemic Versus Epidemic Influenza. *Am J Epidemiol*. **173**, 127–135 (2011).
48. A. Colas de la Noue, M. Estienney, S. Aho, J.-M. Perrier-Cornet, A. de Rougemont, P. Pothier, P. Gervais, G. Belliot, Absolute Humidity Influences the Seasonal Persistence and Infectivity of Human Norovirus. *Appl Environ Microbiol*. **80**, 7196–7205 (2014).

49. H. Qiu, J. Wu, L. Hong, Y. Luo, Q. Song, D. Chen, Clinical and epidemiological features of 36 children with coronavirus disease 2019 (COVID-19) in Zhejiang, China: an observational cohort study. *The Lancet Infectious Diseases*. **20**, 689–696 (2020).
50. A. A. Kelvin, S. Halperin, COVID-19 in children: the link in the transmission chain. *The Lancet Infectious Diseases*. **20**, 633–634 (2020).
51. R. M. Viner, S. J. Russell, H. Croker, J. Packer, J. Ward, C. Stansfield, O. Mytton, C. Bonell, R. Booy, School closure and management practices during coronavirus outbreaks including COVID-19: a rapid systematic review. *The Lancet Child & Adolescent Health*. **4**, 397–404 (2020).
52. R. N. Salas, J. M. Shultz, C. G. Solomon, The Climate Crisis and Covid-19 — A Major Threat to the Pandemic Response. *New England Journal of Medicine*. **383**, e70 (2020).

Table 1. GAM Regression against new cases from March 10, 2020 to September 29, 2020. The standard errors are shown in parenthesis.

	Group 1	Group 2	Group 3	Group 4	Group 5	Group 6	Group 7	Group 8	Group 9	Group 10
<i>Predictors</i>	<i>Estimates</i>	<i>Estimates</i>	<i>Estimates</i>	<i>Estimates</i>	<i>Estimates</i>	<i>Estimates</i>	<i>Estimates</i>	<i>Estimates</i>	<i>Estimates</i>	<i>Estimates</i>
Intercept	-7.292 *** (0.018)	-10.406 *** (0.061)	-12.889 *** (0.213)	-7.991 *** (0.013)	-10.152 *** (0.037)	-9.527 *** (0.030)	-8.223 *** (0.021)	-8.570 *** (0.037)	-9.443 *** (0.028)	-8.168 *** (0.058)
Absolute Humidity [†] (14-day Lag)	-0.114 *** (0.001)	-0.013 *** (0.002)	-0.028 *** (0.002)	-0.242 *** (0.001)	0.022 *** (0.001)	-0.006 *** (0.001)	-0.160 *** (0.001)	-0.011 *** (0.001)	-0.119 *** (0.001)	-0.006 ** (0.002)
Non-essential Visitations (14-day Lag)	-0.090 *** (0.012)	1.376 *** (0.067)	0.353 *** (0.052)	0.827 *** (0.012)	1.077 *** (0.023)	0.282 *** (0.014)	-0.195 *** (0.016)	-0.364 *** (0.013)	0.895 *** (0.012)	-0.393 *** (0.057)
spline(time)	8.992 ***	8.937 ***	8.982 ***	8.999 ***	8.987 ***	8.995 ***	8.994 ***	8.977 ***	8.998 ***	8.965 ***
Observations	19834	4678	6020	38602	16301	16264	14670	22595	21067	3154
R ²	0.569	0.638	0.625	0.655	0.365	0.629	0.772	0.556	0.675	0.457

* $p < 0.05$ ** $p < 0.01$ *** $p < 0.001$

[†] Negative coefficient estimates are bolded

[‡] Estimated degree of freedom for spline basis function are listed

Table 2. GAM Regression against new cases from March 10, 2020 to June 30, 2020. The standard errors are shown in parenthesis.

	Group 1	Group 2	Group 3	Group 4	Group 5	Group 6	Group 7	Group 8	Group 9	Group 10
<i>Predictors</i>	<i>Estimates</i>	<i>Estimates</i>	<i>Estimates</i>	<i>Estimates</i>	<i>Estimates</i>	<i>Estimates</i>	<i>Estimates</i>	<i>Estimates</i>	<i>Estimates</i>	<i>Estimates</i>
Intercept	-8.398 *** (0.033)	-12.221 *** (0.230)	-10.524 *** (1.008)	-8.902 *** (0.024)	-9.228 *** (0.058)	-12.094 *** (0.060)	-10.061 *** (0.057)	-9.806 *** (0.083)	-11.987 *** (0.036)	-7.669 *** (0.096)
Absolute Humidity [†] (14-day Lag)	-0.123 *** (0.002)	-0.017 ** (0.006)	-0.000 (0.006)	-0.254 *** (0.001)	-0.060 *** (0.002)	0.054 *** (0.002)	-0.194 *** (0.002)	-0.017 *** (0.001)	0.035 *** (0.002)	0.034 *** (0.007)
Non-essential Visitations (14-day Lag)	0.854 *** (0.035)	1.422 *** (0.128)	-0.614 *** (0.100)	0.993 *** (0.024)	0.976 *** (0.036)	0.585 *** (0.029)	-0.223 *** (0.035)	-0.183 *** (0.029)	3.110 *** (0.027)	-1.849 *** (0.115)
spline(time) [‡]	8.989 ***	8.924 ***	8.821 ***	8.996 ***	8.974 ***	8.981 ***	8.971 ***	8.942 ***	8.996 ***	8.645 ***
Observations	8870	2103	2537	18573	7622	7852	6630	10532	9986	1445
R ²	0.651	0.583	0.743	0.652	0.601	0.713	0.749	0.664	0.756	0.473

* $p < 0.05$ ** $p < 0.01$ *** $p < 0.001$

[†] Negative coefficient estimates are bolded

[‡] Estimated degree of freedom for spline basis function are listed

0

1

Table 3. GAM Regression against new cases from July 1, 2020 to September 29, 2020. The standard errors are shown in parenthesis.

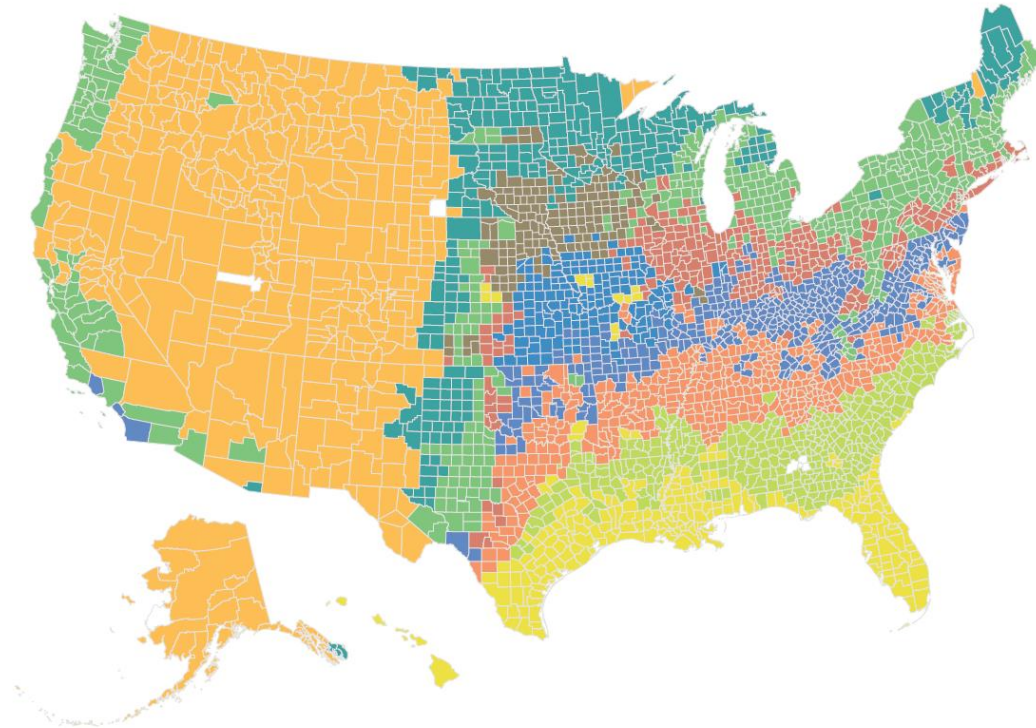
	Group 1	Group 2	Group 3	Group 4	Group 5	Group 6	Group 7	Group 8	Group 9	Group 10
<i>Predictors</i>	<i>Estimates</i>	<i>Estimates</i>	<i>Estimates</i>	<i>Estimates</i>	<i>Estimates</i>	<i>Estimates</i>	<i>Estimates</i>	<i>Estimates</i>	<i>Estimates</i>	<i>Estimates</i>
Intercept	-10.444 *** (0.046)	-12.546 *** (0.094)	-14.296 *** (0.266)	-9.361 *** (0.028)	-12.846 *** (0.058)	-8.502 *** (0.050)	-7.736 *** (0.034)	-10.581 *** (0.065)	-11.816 *** (0.073)	-12.645 *** (0.126)
Absolute Humidity† (14-day Lag)	-0.007 *** (0.001)	0.023 *** (0.003)	-0.023 *** (0.003)	-0.095 *** (0.001)	0.114 *** (0.002)	-0.051 *** (0.001)	-0.164 *** (0.001)	0.012 *** (0.001)	-0.026 *** (0.002)	0.016 *** (0.003)
Non-essential Visitations (14-day Lag)	2.060 *** (0.040)	4.541 *** (0.101)	1.463 *** (0.115)	2.183 *** (0.032)	2.329 *** (0.035)	0.674 *** (0.022)	0.350 *** (0.036)	1.427 *** (0.038)	1.555 *** (0.044)	3.790 *** (0.112)
spline(time)‡	8.803 ***	8.955 ***	8.854 ***	8.970 ***	8.987 ***	8.997 ***	8.970 ***	8.915 ***	8.909 ***	8.947 ***
Observations	10964	2575	3483	20029	8679	8412	8040	12063	11081	1709
R ²	0.650	0.647	0.607	0.727	0.352	0.619	0.792	0.541	0.816	0.556

* $p < 0.05$ ** $p < 0.01$ *** $p < 0.001$

† Negative coefficient estimates are bolded

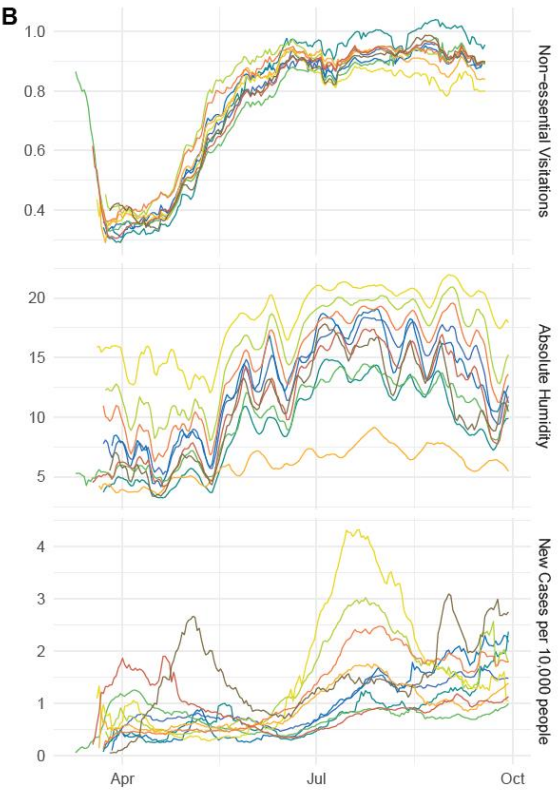
‡ Estimated degree of freedom for spline basis function are listed

A



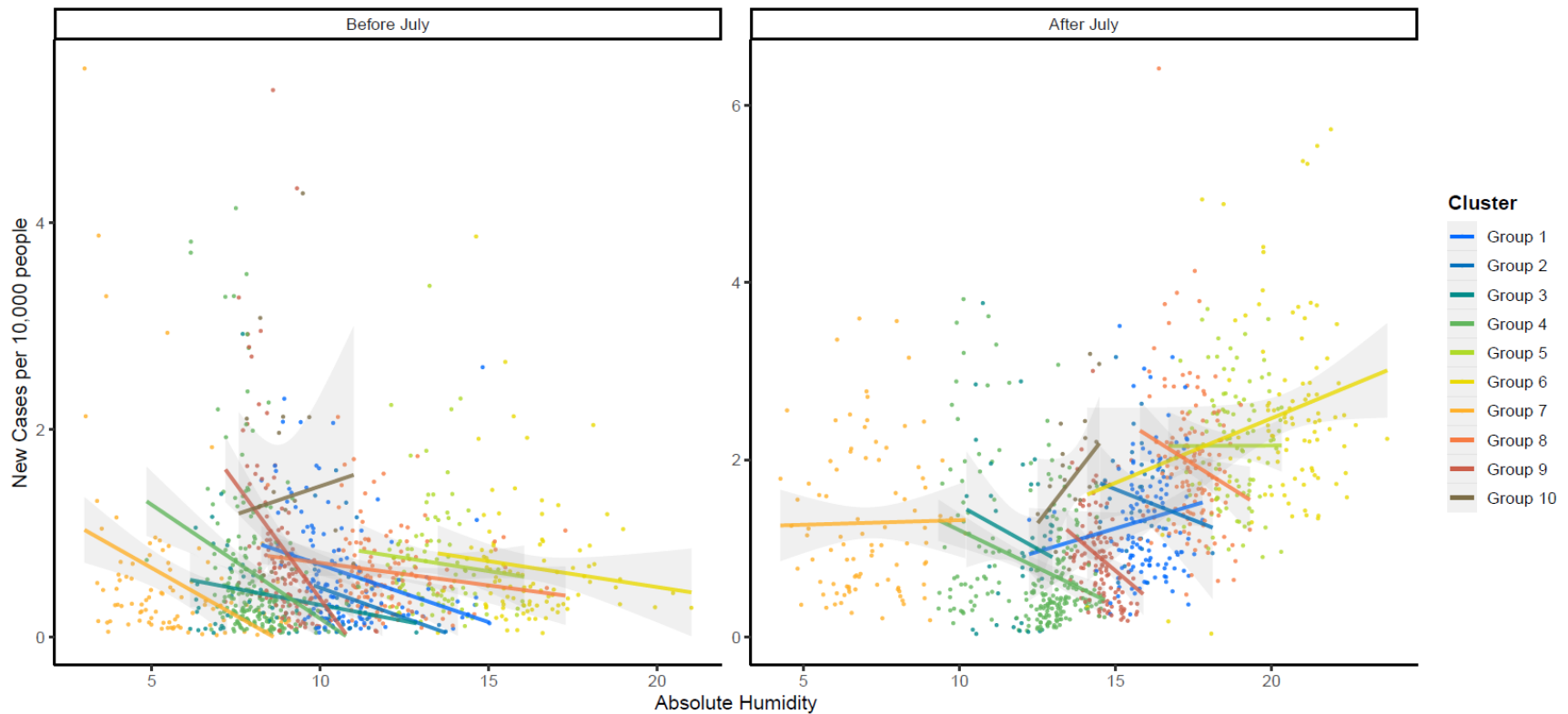
Cluster ■ Group 1 ■ Group 2 ■ Group 3 ■ Group 4 ■ Group 5 ■ Group 6 ■ Group 7 ■ Group 8 ■ Group 9 ■ Group 10 ■ NA

B



4

5 **Figure 1.** (A) Map of US Counties and their respective absolute humidity clusters. The clustering analysis was conducted using a
 6 partitional algorithm that utilized dynamic time warping (DTW) to measure similarity between absolute humidity profiles of 3,137
 7 counties in the United States. Expectantly, the clustering of absolute humidity is related to the geography of the counties which serves
 8 as a proxy for regional weather patterns and different climatological regimes. (B) The cross-sectional smoothed mean of human
 9 encounter absolute humidity, and new case per 10,000 people trends for each cluster group of the 987 counties analyzed in the
 10 regression analysis.



11

12 **Figure 2.** The average daily new cases plotted against average absolute humidity of 987 counties for March 10 to June 30 (*left*) and
 13 July 1 to September 29 (*right*). For each cluster group, we added a simple linear trend line with shaded standard errors. We see that
 14 within each humidity cluster before July, the absolute humidity in regions with lower humidity, i.e. Groups 6 and 7, has a larger
 15 negative slope correlation; however, after July when humidity patterns are stable, this relationship is less noticeable without
 16 considering the temporal relationships. The wetter regions also experience less impact from absolute humidity.