

Weather variables impact on COVID-19 incidence

Javier G. Corripio^{1*}, Lorna Raso¹

¹ meteoexploration.com, Innsbruck, Austria

* jgc@meteoexploration.com

Abstract

We test the hypothesis of COVID-19 contagion being influenced by meteorological parameters such as temperature or humidity. We analysed data at high spatial resolution (regions in Italy and counties in the USA) and found that while at low resolution this might seem the case, at higher resolution no correlation is found. Our results are consistent with a poor outdoors transmission of the disease. However, a possible indirect correlation between good weather and a decrease in disease spread may occur, as people spend longer time outdoors.

Introduction

Weather impact on the transmission of SARS-CoV-2 virus causing the current COVID-19 pandemic is of interest to determine whether the infection might be seasonal. This could help contagion forecast models and be useful to manage health care resources. It is an area of active current research. Promising results were published at the initial stages of disease dispersion, showing graphically the correlation between the disease outbreak and narrow bands of air temperature or absolute humidity [1] as shown in Fig. 1. If the relationship between weather parameters and disease transmission could be established in a clear way, it would be a useful forecasting tool. Our aim is to evaluate if the relationship is maintained for a longer period of time, on a wider range of weather conditions and at a higher spatial resolution.

The relationship between meteorological parameters and disease spread has been investigated more amply in relation to the influenza outbreaks. In temperate regions influenza outbreaks show marked seasonality with an onset in autumn that continues through the winter. The cause of this seasonality is not fully understood. Travel [2], absolute humidity [3,4] or school calendar [5] all appear to affect the spread of influenza, but the mechanisms behind these relationships are still unknown. Whereas other respiratory viruses seem to lack seasonality or peak in spring or summer, coronavirus outbreaks tend to begin in winter and end before the summer [6,7].

Our intention is to find a simple meteorological model that can indicate whether meteorological parameters have any impact on transmission rates. We analyse the incidence data at high spatial resolution, for 21 regions in Italy and for 3142 counties in the USA, with weather data at comparable resolution interpolated to the center of every region or county.

Materials and methods

Meteorological data

Meteorological data were obtained from the NCEP GDAS/FNL 0.25 Degree Global Tropospheric Analyses and Forecast Grids [8] and the hourly ERA5 reanalysis from the European Centre for Medium Range Weather Forecasts ECMWF [9]. We compared both datasets to check if discrepancies are high. On average there is less than one degree centigrade difference for the daily average temperature and less than 1g/kg in specific humidity. As the ERA5 has higher temporal resolution, we used this data set for the statistical analysis.

Both ERA5 and GDAS provide the data in grib format (GRIdded Binary), that was processed using the Grid Analysis and Display System (GrADS) software [10]. Data were interpolated for every county and region coordinates from the beginning of January to the seventh of April, two weeks after lockdown was declared in many US states. This time frame was used so that both in Italy and the USA the window between the first cases and lockdown was analysed. In this way it is possible to compare spread before any lockdown effects. Italy went into lockdown on the ninth of March and USA between the 19th and the 24th of March. The resulting csv files with meteorological data are available with the supplementary material at <https://meteoexploration.com/covid19/>.

ERA5 does not provide specific or absolute humidity, but dew point temperature. From this variable and the local pressure we derived first water vapor pressure, following Bolton (Eq. 10) [11], and then specific and absolute humidity following Stull (Eq. 4.7 and 4.10) [12].

$$e_0 = 6.112 e^{((17.67 T_d)/(T_d+243.5))} \quad (1)$$

$$q = \frac{0.622 e_0}{P - (0.378 e_0)} \quad (2)$$

$$\rho_v = \frac{100 e_0}{R_v T_k}, \quad (3)$$

where e_0 is actual water vapour pressure, T_d is dew point temperature in degrees centigrade, q is specific humidity in kg/kg, P is local atmospheric pressure in hPa, $0.622 = 18.016/28.966$ is the ratio of molecular weights of water and dry air, ρ_v is absolute humidity, R_v is the gas constant for water vapor ($461 JK^{-1}kg^{-1}$) and T_k is air temperature in Kelvin.

Different studies use either specific or absolute humidity, but the differences in the range of temperature and pressures found in this study are small. For the whole of Italy, for example, the mean difference is smaller than 1g of water vapour and values run almost parallel to each other.

Hourly data was aggregated as daily means and later aggregated as the mean over a period of eight days with a lag of two days with respect to incidence cases. The median incubation period of COVID-19 has been estimated to be 5.1 days, with 97.5% of those developing symptoms doing so within 11.5 days [13]. Thus, we tested a range of days and lags from seven to 14 and two to five respectively. We imposed a minimum of seven days for the accumulated incidence to avoid weekend errors in the data reporting. The final figures of eight and two gave the best result in the regression analysis.

Epidemiological data

We retrieved global data on cases and mortality from the COVID-19 Data Repository by the Center for Systems Science and Engineering (CSSE) at Johns Hopkins University [14] and from the European Centre for Disease Prevention and Control [15]. Both datasets are consistent and updated daily. Italian data at regional level were downloaded from the github repository of the *Sito del Dipartimento della Protezione Civile* [16]. Data for the USA at county level were also obtained from the Johns Hopkins University repository. An additional attempt was made to use data from Spain, a country that has been heavily impacted by the pandemic. There are regional data for Spain, but this is too coarse a resolution, a single region such as Andalucía is larger than several European countries. Provincial data are very fragmented, without a single central repository. A crowdsourced effort to compile these data [17] has been made, but unfortunately the result has too many gaps.

Geographical and other data

Global and European shapefiles are from Natural Earth [18]. Initially we used the taxpayer funded Eurostat GISCO: Geographical Information and maps - NUTS, but data usage is too restrictive while Natural Earth is completely open. Shapefiles for the USA were downloaded from the TIGER/Line Shapefiles from the United States Census Bureau [19]. Data for the population of Italian regions were obtained from the Italian Istat - *Istituto nazionale di statistica* [20]. For US county population, data were obtained from the United States Census Bureau [21]. Additional, non conclusive tests were done with airport transportation data, that were obtained from the EU Open Data Portal [22] and the Federal Aviation Administration [23].

Statistical analysis

A first step is using descriptive statistics and spatial visualisation that can help identifying any possible correlation between weather and disease transmission. A second step is designed following Barreca and Shimshack [24]. They found that half of the average seasonal differences in US influenza mortality can be explained by seasonal differences in humidity alone, and there seems to be an important threshold at 6g/kg. Because mortality is dependant not only on contagion, but also on the health system capacity and resources, we used the number of cases accumulated over a period of time. Following Barreca and Shimshack we assume that incidence can be simulated with a simple statistical model as:

$$I = \sum_{i=1}^{i=n} f_i(x_i) + \epsilon \quad (4)$$

where $f_i(x_i)$ is a function of variable x_i and ϵ a general error term.

This model was implemented in the R statistical package using generalised linear models (glm) [25] and generalised additive models (gam) from the *mgcv* library [26]. We are interested in response to meteorological variables, but we also tested population density and passenger emplanment at nearest airport with very poor results.

Results and discussion

The visualisation of Italian data in Figures 2 to 4 seems to confirm the hypothesis that transmission is hindered by specific humidity above 6g/kg and mean air temperature

above 11°C [1]. Fig. 2 shows the evolution of accumulated incidence during eight days in Italy from the beginning of February to the 23rd of March, two weeks after general lockdown. Every colour represents one region and the points increase in size with the date, the name of the region is plotted at the last date. It appears that regions with specific humidity higher than 6 g/kg of water vapour at the beginning of the period did not develop many cases. Similarly in Fig. 3, a threshold slightly over 11°C seems to hinder the development of cases. Fig. 4 shows the mean ultraviolet radiation, and it appears that regions with values higher than 18 Wm⁻² at the beginning of the episode remain at very low incidence. Valle d'Aosta seems to be an exception, but the increase in UV radiation is towards the end, while at the beginning of the period values were below the apparent threshold. This is a small region at high altitude in the Alps, bordering with Switzerland, where clear days in March brought about a high dose of UV radiation.

Additional tests were done for precipitation, but as Fig. 5 shows there is no evident relationship, thus we discarded this variable for further analysis. A statistical model using generalised linear models with Poisson error structure indicate that all three parameters (specific humidity, temperature and ultraviolet radiation) are significant, but the results have a very low value of adjusted R-squared, as low as 0.14. A model with generalised additive models seems to explain 32% of the deviance, with a poor R-squared value of 0.18. We used count data, incidence accumulated over eight days for the models but the figures show incidence per 100 000 inhabitants to help the visualisation.

The threshold hypothesis, which could be useful in forecasting future seasonal incidence, is refuted by the high resolution data from the USA. As can be noted from Figures 6 to 8, there is no evidence of thresholds whatsoever in these data. Fig. 6 shows incidence per capita with respect to Specific Humidity. There are too many counties to use distinctive colours, so these are depicted as a range from grey to blue according to the initial values. Although there seems to be some clustering, especially at high values, there is a broad dispersion at lower values for most of the points. The range of the dispersion is limited by the maximum mean temperature at every county at this period of the year, as maximum possible specific humidity increases exponentially with temperature. Blaine, New York, Rockland and Westchester counties are outside the chart. The statistical modelling with glm results in a very low R squared value of 0.02 and only temperature seems to be significant, but this is probably a spurious correlation, as the number of cases increase from winter towards the spring, which correspond to a seasonal increase in temperature.

Our results suggest that the spread of SARS-CoV-2 is not influenced by outdoors specific humidity, temperature or UV radiation. If we consider some of the literature on influenza spread, several experiments show that in closed spaces increased humidity reduces the influenza virus droplet transmissions [27–29]. The mechanics behind this effect has been explored by Yang and Marr [30]. Kormuth et al. [31] found this is probably true only for droplet transmission as they confirmed aerosol transmission to be the same even at high relative humidity. However, there does not seem to be much literature on contagion rates in the open air.

Jaakkola et al. [32] found that changes in temperature and absolute humidity were more likely to cause infection in Finnish military conscripts. They suggest this may be due to cold stress and a change in the susceptibility of the air passage to viral infections. This is consistent with Davey and Ried [33] who also found an increase in influenza cases with a decrease in temperature in the UK, a finding supported by Davies et al. [34] who correlated increased pulmonary disease mortality rates with unusually cold and dry weather spells. Willem et al. [35] find that part of this weather effect may be behavioural. As with warmer temperatures people spend more time outside, therefore

decreasing the contact with people in closed spaces. This may go some way to explaining the differences between USA and Italy. In Italy much of the spread was late winter whereas in the USA it was already the beginning of Spring, when days are getting longer and people spend more time outside. It is also a plausible explanation for Italy, where most of the cases were concentrated in the Po Valley. The Po is often cold and cloudy from autumn to spring, a weather less inviting to outdoor life than in the southern regions, where clear skies and moderate cold are common even in winter. A poor outdoors contagion is also evidenced by the high rate of contagion in some ski resorts of the Alps, such As Valle d'Aosta in Italy or Ischgl and Sölden in Tirol, where most of the cases can be traced back to just a few après ski centres [36].

Conclusion

Although initial data at coarse resolution seems to suggest a narrow range of meteorological conditions for the spread of COVID-19, a closer look at higher resolution does not support this hypothesis. Higher resolution data reveal no correlation between meteorological parameters and COVID-19 incidence. This lack of correlation may be indicative of a poor outdoor transmission. It could be that lower outdoor transmissions have an indirect relationship with meteorological parameters. The good weather and longer days, will affect behaviour as people tend to spend more time outdoors, making the spread of the disease less likely. It may also explain the differences in contagion rates between the north and the south of Italy, due mostly to behavioural differences, and it is supported by the high relative importance of a few “hot spots” in the spreading of the disease, in agreement with Endo’s work [37].

Supporting information

S1 Data availability. Meteorological data interpolated to the coordinates of every region in Italy and every county in the USA, figures at higher resolution and additional figures are available at <https://meteoexploration.com/covid19/>.

Acknowledgments

We are grateful to the different Agencies, Research Centres and Universities that have made data readily available. meteoexploration.com provided computer resources and staff time. We did not receive any funding for this work.

References

1. Sajadi MM, Habibzadeh P, Vintzileos A, Shokouhi S, Miralles-Wilhelm F, Amoroso A. Temperature and Latitude Analysis to Predict Potential Spread and Seasonality for COVID-19. SSRN Electronic Journal. 2020;(410):6–7. doi:10.2139/ssrn.3550308.
2. Brownstein J, Wolfe C, Mandl K. Empirical Evidence for the Effect of Airline Travel on Inter-Regional Influenza Spread in the United States. PLoS Med. 2006;3:e401. doi:10.1371/journal.pmed.0030401.
3. Shaman J, Kohn M. Absolute humidity modulates influenza survival, transmission, and seasonality. Proceedings of the National Academy of Sciences

- of the United States of America. 2009;106(9):3243–3248.
doi:10.1073/pnas.0806852106.
4. Shaman J, Pitzer VE, Viboud C, Grenfell BT, Lipsitch M. Absolute humidity and the seasonal onset of influenza in the continental United States. *PLoS Biology*. 2010;8(2). doi:10.1371/journal.pbio.1000316.
 5. Chao DL, Halloran ME, Lonfini Jr IM. School opening dates predict pandemic influenza A (H1N1) epidemics in the USA. *J Infect Dis*. 2010;202:877–880. doi:10.1086/655810.
 6. Moriyama M, Hugentohler WJ, Iwasaki A. Seasonality of Respiratory Viral Infections. *Annual Review of Virology*. 2020;7:2.1–2.19. doi:10.1146/annurev-virology-012420-022445.
 7. Morikawa S, Kohdera U, Hosaka T, Ishii K, Akagawa S, Hiroi S, et al. Seasonal variations of respiratory viruses and etiology of human rhinovirus infection in children. *Journal of Clinical Virology*. 2015;73:14–19. doi:10.1016/j.jcv.2015.10.001.
 8. NCEP GDAS/FNL 0.25 Degree Global Tropospheric Analyses and Forecast Grids; 2015. Available from: <https://doi.org/10.5065/D65Q4T4Z>.
 9. ERA5 Reanalysis;. Available from: <https://cds.climate.copernicus.eu/cdsapp#!/home>.
 10. Grid Analysis and Display System (GrADS);. Available from: <http://cola.gmu.edu/grads/>.
 11. Bolton D. The computation of equivalent potential temperature. *Monthly Weather Review*. 1980;108:1046–1053.
 12. Stull R. *Practical Meteorology: An Algebra-based Survey of Atmospheric Science*. Vancouver: Univ. of British Columbia; 2017. Available from: https://www.eoas.ubc.ca/books/Practical_Meteorology/prmet102/Practical_Meteorology-v1.02b-WholeBookColor.pdf.
 13. Lauer SA, Grantz KH, Bi Q, Jones FK, Zheng Q, Meredith HR, et al. The Incubation Period of Coronavirus Disease 2019 (COVID-19) From Publicly Reported Confirmed Cases: Estimation and Application. *Annals of internal medicine*. 2020;2019. doi:10.7326/M20-0504.
 14. COVID-19 Data Repository by the Center for Systems Science and Engineering (CSSE) at Johns Hopkins University;. Available from: <https://github.com/CSSEGISandData/COVID-19>.
 15. European Center for Disease Prevention and Control (ECDC);. Available from: <https://www.ecdc.europa.eu/en/covid-19-pandemic>.
 16. Sito del Dipartimento della Protezione Civile - Emergenza Coronavirus: la risposta nazionale;. Available from: <https://github.com/pcm-dpc/COVID-19>.
 17. Propagación del COVID-19 por regiones y provincias en diferentes países;. Available from: <https://lab.montera34.com/covid19/datos.html>.
 18. Natural Earth Data;. Available from: <https://www.naturalearthdata.com>.

19. TIGER/Line Shapefiles - United States Census Bureau;. Available from: <https://www.census.gov/geographies/mapping-files/time-series/geo/tiger-line-file.html>.
20. Istat - Istituto nazionale di statistica;. Available from: <http://dati.istat.it/Index.aspx?DataSetCode=DCIS{ }POPRES1>.
21. United States Census Bureau;. Available from: <https://www.census.gov/data/datasets/time-series/demo/popest/2010s-counties-total.html>.
22. EU Open Data Portal;. Available from: <https://data.europa.eu/euodp/en/data/dataset/43C6uGqWp92dX7vlgNzJA>.
23. Federal Aviation Administration;. Available from: <https://www.faa.gov/airports/planning{ }capacity/passenger{ }allcargo{ }stats/passenger/>.
24. Barreca AI, Shimshack JP. Absolute humidity, temperature, and influenza mortality: 30 years of county-level evidence from the united states. *American Journal of Epidemiology*. 2012;176(SUPPL. 7):114–122. doi:10.1093/aje/kws259.
25. Crawley MJ. *The R book*. 2nd ed. Chichester: Wiley; 2012.
26. Wood SN. *Generalized Additive Models. An Introduction with R*. 2nd ed. New York: Chapman and Hall/CRC; 2017.
27. Lowen AC, Steel J. Roles of Humidity and Temperature in Shaping Influenza Seasonality. *Journal of Virology*. 2014;88(14):7692–7695. doi:10.1128/jvi.03544-13.
28. Steel J, Palese P, Lowen AC. Transmission of a 2009 Pandemic Influenza Virus Shows a Sensitivity to Temperature and Humidity Similar to That of an H3N2 Seasonal Strain. *Journal of Virology*. 2011;85(3):1400–1402. doi:10.1128/jvi.02186-10.
29. Noti JD, Blachere FM, McMillen CM, Lindsley WG, Kashon ML, Slaughter DR, et al. High Humidity Leads to Loss of Infectious Influenza Virus from Simulated Coughs. *PLoS ONE*. 2013;8(2):2–9. doi:10.1371/journal.pone.0057485.
30. Yang W, Marr LC. Mechanisms by which ambient humidity may affect viruses in aerosols. *Applied and Environmental Microbiology*. 2012;78(19):6781–6788. doi:10.1128/AEM.01658-12.
31. Kormuth KA, Lin K, Prussin AJ, Vejerano EP, Tiwari AJ, Cox SS, et al. Influenza virus infectivity is retained in aerosols and droplets independent of relative humidity. *Journal of Infectious Diseases*. 2018;218(5):739–747. doi:10.1093/infdis/jiy221.
32. Jaakkola K, Saukkoriipi A, Jokelainen J, Juvonen R, Kauppila J, Vainio O, et al. Decline in temperature and humidity increases the occurrence of influenza in cold climate. *Environmental Health: A Global Access Science Source*. 2014;13(1):1–8. doi:10.1186/1476-069X-13-22.
33. Davey ML, Reid D. Relationship of air temperature to outbreaks of influenza. *British journal of preventive & social medicine*. 1972;26(1):28–32. doi:10.1136/jech.26.1.28.
34. Davis RE, Rossier CE, Enfield KB. The impact of weather on influenza and pneumonia mortality in new york city, 1975-2002: A retrospective study. *PLoS ONE*. 2012;7(3):1–8. doi:10.1371/journal.pone.0034091.

35. Willem L, van Kerckhove K, Chao DL, Hens N, Beutels P. A Nice Day for an Infection? Weather Conditions and Social Contact Patterns Relevant to Influenza Transmission. PLoS ONE. 2012;7(11). doi:10.1371/journal.pone.0048695.
36. Grill P. How an Austrian ski paradise became a COVID-19 hotspot;. Available from: <https://www.euractiv.com/section/coronavirus/news/ischgl-oesterreichisches-skiparadies-als-corona-hotspot/>.
37. Endo A, Abbott S, Kucharski AJ, Funk S. Estimating the overdispersion in COVID-19 transmission using outbreak sizes outside China. Wellcome Open Research. 2020;5:67. doi:10.12688/wellcomeopenres.15842.1.

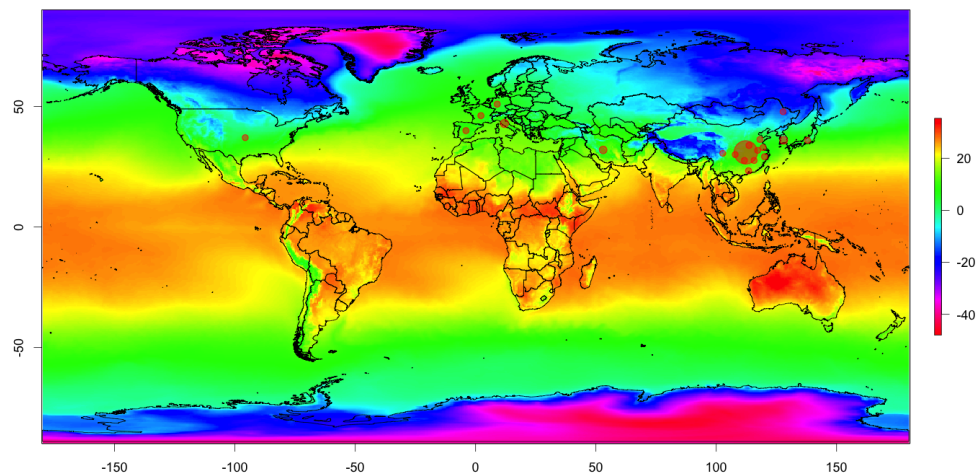


Fig 1. World temperature and cases. World average temperature [°C] during February and location of twenty spots with the higher number of COVID19 cases at the beginning of March, 2020. All spots are within a similar range of mean temperature except for Heilongjiang in northern China.

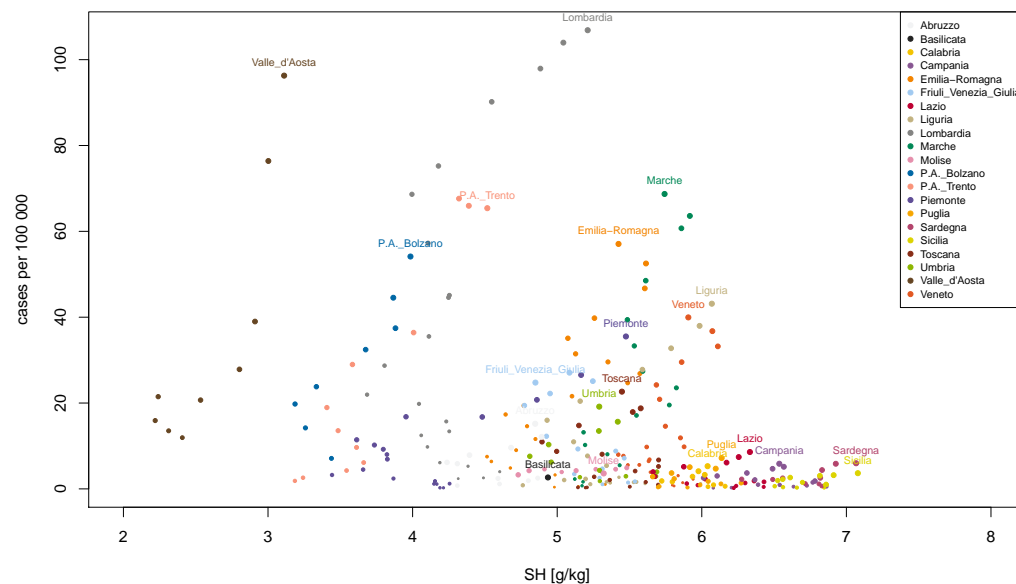


Fig 2. Evolution of cases in Italy and specific humidity. Cases per region in Italy against specific humidity. Every dot colour corresponds to an Italian region, size of dots increase with time from the 11th of February to the 23rd of March, two weeks after Italy's lockdown.

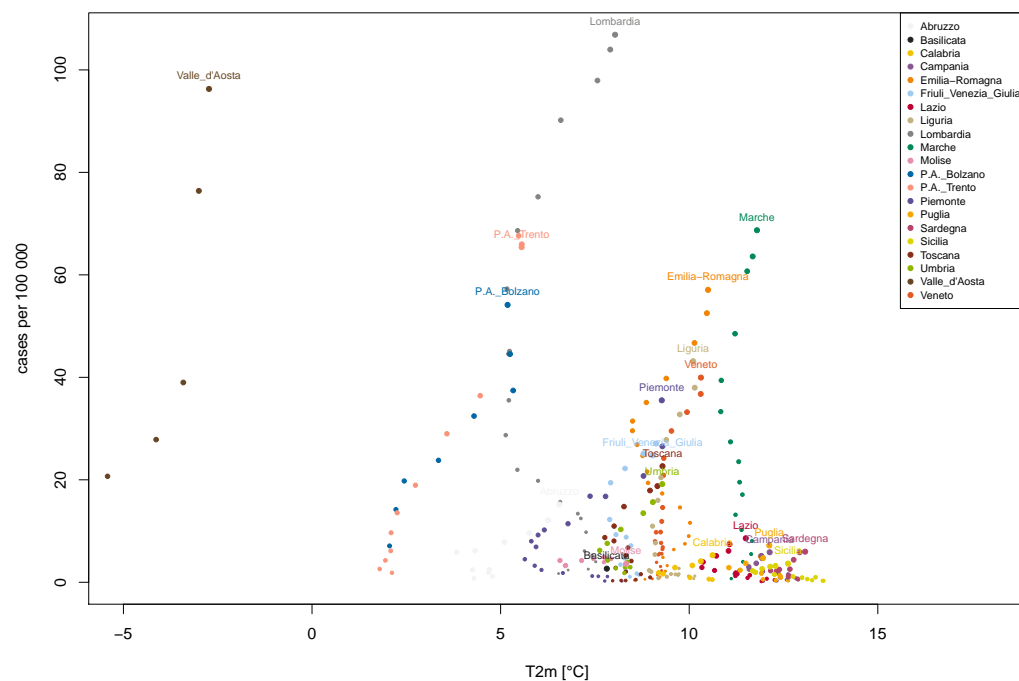


Fig 3. Evolution of cases in Italy and temperature. Cases per region in Italy against air temperature. Every dot colour corresponds to an Italian region, size of dots increase with time from the 11th of February to the 23rd of March, two weeks after Italy's lockdown.

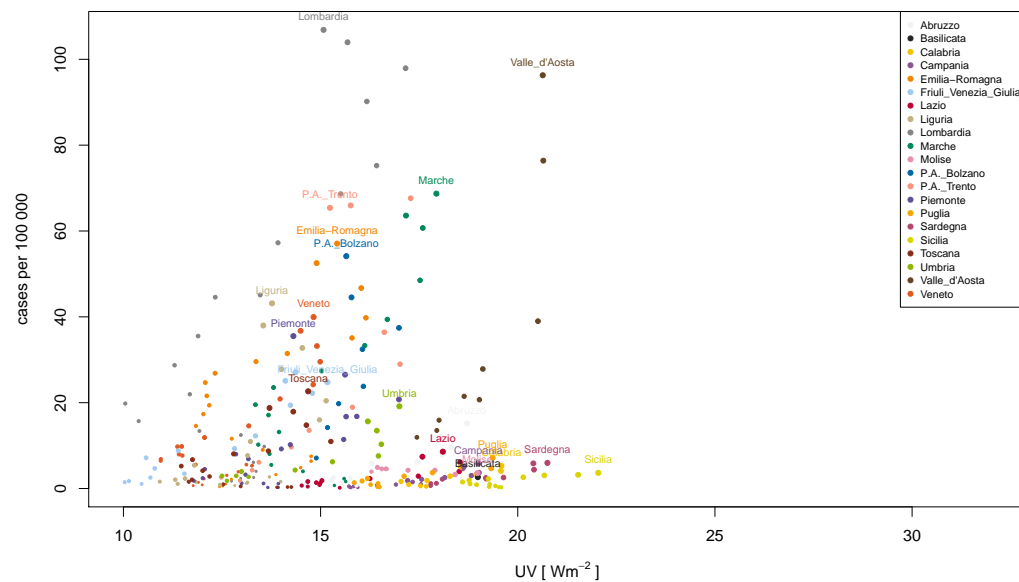


Fig 4. Evolution of cases in Italy and UV radiation. Cases per region in Italy against ultraviolet radiation. Every dot colour corresponds to an Italian region, size of dots increase with time from the 11th of February to the 23rd of March, two weeks after Italy's lockdown.

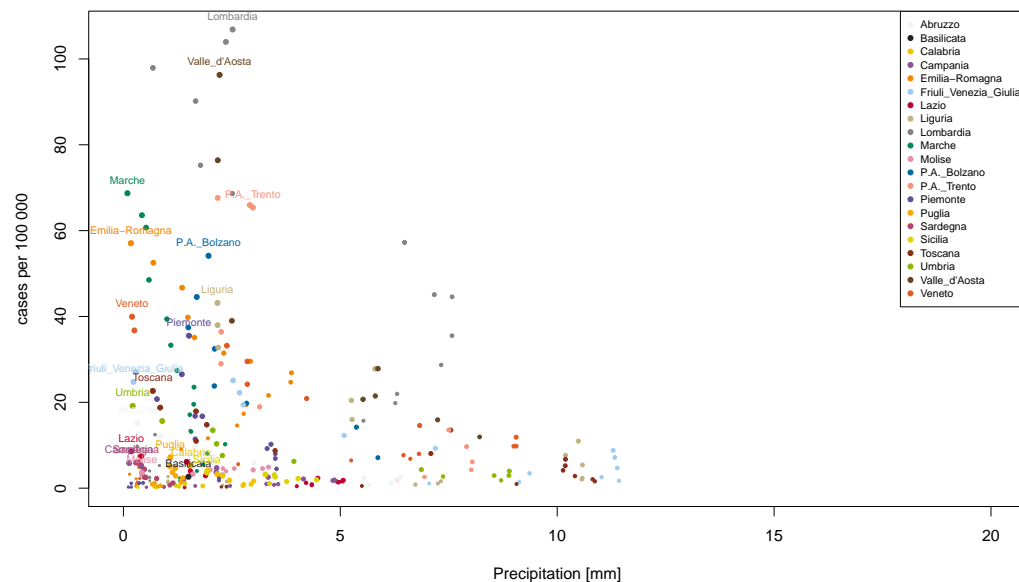


Fig 5. Evolution of cases in Italy and precipitation. Cases per region in Italy against precipitation. Every dot colour corresponds to an Italian region, size of dots increase with time from the 11th of February to the 23rd of March, two weeks after Italy's lockdown.

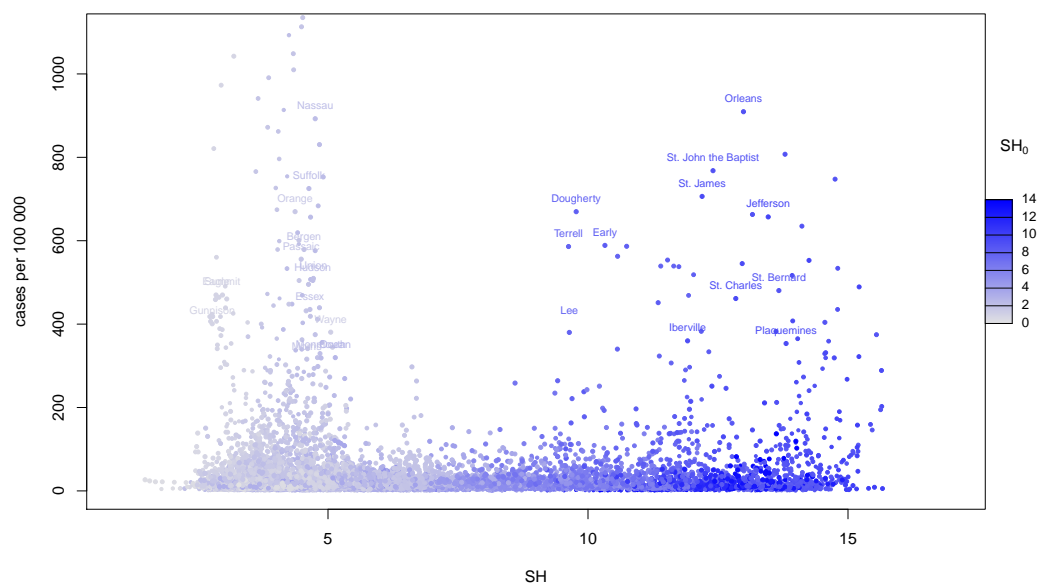


Fig 6. Evolution of cases in the USA and specific humidity. Cases per county in the USA against specific humidity. Dot colours are graded from grey to blue according to initial values of specific humidity (low to high), size of dots increase with time from the 20th of February to the seventh of April, between 10 days to two weeks after lockdown depending on State.

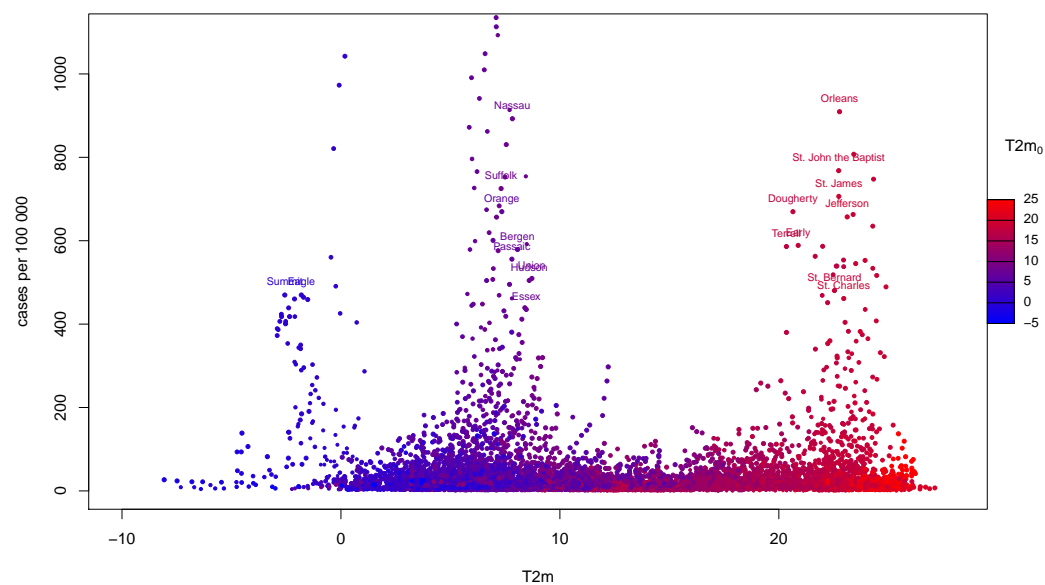


Fig 7. Evolution of cases in the USA and temperature. Cases per county in the USA against air temperature. Dot colours are graded from blue to red according to initial values of air temperature (low to high), size of dots increase with time from the 20th of February to the seventh of April, between 10 days to two weeks after lockdown depending on State.

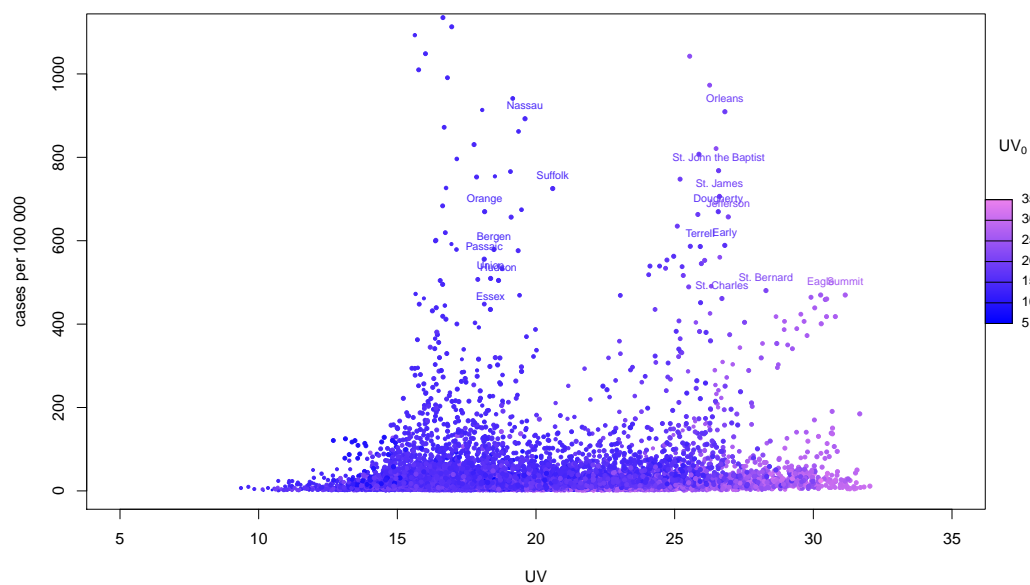


Fig 8. Evolution of cases in the USA and UV. Cases per county in the USA against ultraviolet radiation. Dot colours are graded from blue to violet according to initial values of ultraviolet radiation (low to high), size of dots increase with time from the 20th of February to the seventh of April, between 10 days to two weeks after lockdown depending on State.