

LETTER • OPEN ACCESS

Weather drives variation in COVID-19 transmission and detection

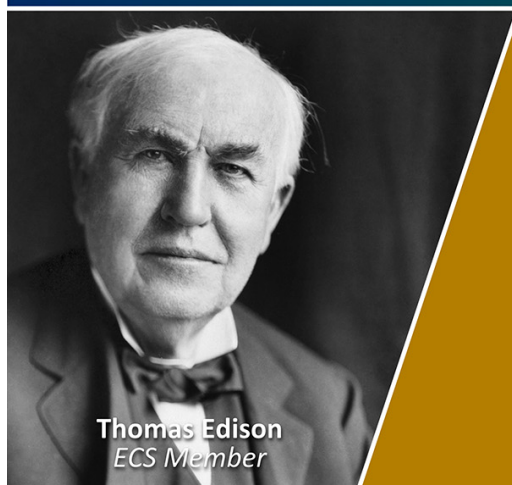
To cite this article: James Rising *et al* 2023 *Environ. Res.: Climate* **2** 011001

View the [article online](#) for updates and enhancements.

You may also like

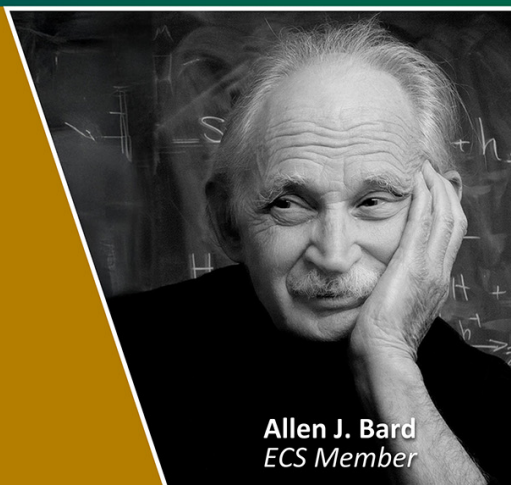
- [The Impact of Weather on Traffic Speed in Urban Area](#)
Jacek Chmielewski and Marcin Budzynski
- [How will climate change shape climate opinion?](#)
Peter D Howe, Jennifer R Marlon, Matto Mildenberger et al.
- [The impact of weather changes on air quality and health in the United States in 1994–2012](#)
Iny Jhun, Brent A Coull, Joel Schwartz et al.

Join the Society
Led by Scientists,
for *Scientists Like You!*



The
Electrochemical
Society

Advancing solid state &
electrochemical science & technology



ENVIRONMENTAL RESEARCH CLIMATE



LETTER

Weather drives variation in COVID-19 transmission and detection

OPEN ACCESS

RECEIVED

14 July 2022

REVISED

23 December 2022

ACCEPTED FOR PUBLICATION

13 January 2023

PUBLISHED

26 January 2023

Original Content from
this work may be used
under the terms of the
[Creative Commons
Attribution 4.0 licence](#).

Any further distribution
of this work must
maintain attribution to
the author(s) and the title
of the work, journal
citation and DOI.



James Rising^{1,*} , Manuel Linsenmeier^{2,3} and Ana De Menezes³

¹ School of Marine Science Policy, University of Delaware, Newark, DE, United States of America

² Climate School, Columbia University, New York, NY, United States of America

³ Department of Geography and Environment, London School of Economics, London, United Kingdom

* Author to whom any correspondence should be addressed.

E-mail: jrising@udel.edu

Keywords: COVID-19, temperature, seasonal cycles, Bayesian modeling, transmission, detection

Supplementary material for this article is available [online](#)

Abstract

The debate over the influence of weather on COVID-19 epidemiological dynamics remains unsettled as multiple factors are conflated, including viral biology, transmission through social interaction, and the probability of disease detection. Here we distinguish the distinct dynamics of weather on detection and transmission with a multi-method approach combining econometric techniques with epidemiological models, including an extension of a susceptible-exposed-infectious-recovered model, to analyse data for over 4000 geographic units throughout the year 2020. We find distinct and significant effects of temperature, thermal comfort, solar radiation, and precipitation on the growth of infections. We also find that weather affects the rates of both disease transmission and detection. When we isolate transmission effects to understand the potential for seasonal shifts, the instantaneous effects of weather are small, with R_0 about 0.007 higher in winter than in summer. However, the effects of weather compound over time, so that a region with a 5 °C drop over three months in winter is expected to have 190% more confirmed cases at the end of that 90 days period, relative to a scenario with constant temperature. We also find that the contribution of weather produces the largest effects in high-latitude countries. As the COVID-19 pandemic continues to evolve and risks becoming endemic, these seasonal dynamics may play a crucial role for health policy.

1. Introduction

The debate over the effect of weather on the spread of COVID-19 is still unsettled. In laboratory studies, the survival rate of other coronavirus-related infections, such as MERS-CoV and SARS-CoV, has been confirmed as affected by environmental conditions, particularly temperature and humidity (e.g. Chan *et al* (2011), Leclercq *et al* (2014)). Because SARS-CoV-2, responsible for COVID-19, is expected to mirror such environment-transmission dynamics, several experimental studies have modelled the relationship between COVID-19 and weather. This effort, however, has resulted in inconsistent weather outcomes of interest and lack of consensus of their relative importance to the spread of the disease (SI table 2). For example, various studies suggest high sensitivity of the virus to temperature, humidity and solar radiation (Chin *et al* 2020, Mecenas *et al* 2020, Seyer and Sanlidag 2020, Ganegoda *et al* 2021, Landier *et al* 2021, Majumder and Ray 2021). Yet, some of them suggest that temperature has a significant and negative effect on COVID-19 case growth (Bukhari *et al* 2020, Huang *et al* 2020, Meyer *et al* 2020, Wu *et al* 2020, Ganegoda *et al* 2021, Wilson 2021), while others found that temperature is not as important as other environmental factors, such as radiation (Carleton *et al* 2020, Walrand 2021, Xu *et al* 2021, Yechezkel *et al* 2021). In addition, besides the biology-environment dynamics, individual mobility and social interactions are likely responsive to weather which can, in turn, affect the potential for transmission and detection of the disease (Wang *et al* 2022). For instance, temperature has been found to be positively associated with human mobility, outdoor activities and

travel decisions (Cools *et al* 2010, Nouvellet *et al* 2021, Shao *et al* 2021). Such environmentally driven social behaviour could affect the seasonal cycle of the disease in a specific location as well as generate geographical differences in outbreaks between locations (Carlson *et al* 2020). These complex channels have contributed to our current poor understanding and low ability to model the effect of weather on COVID-19 epidemiological dynamics.

In this study, we use a multi-method approach to examine whether and how weather affects the epidemiological dynamic of COVID-19. In the first part of the analysis, we use a panel regression model with fixed effects (e.g. locality-specific and date-specific intercepts) and exploit the fact that random variation in weather provides a natural experiment to identify the effect of weather on the infectious population and confirmed cases. To isolate two major channels of these estimated effects, specifically transmission and detection, we calibrate an environmentally extended susceptible-exposed-infectious-recovered (SEIR) model, informed by our empirical estimates. In addition, we use this model for a set of idealised simulations which allow us to explain further findings of the empirical analysis.

Our work contributes to the previous literature in several ways. Our main contribution is to demonstrate the potential seasonal component of COVID-19 dynamics using day-to-day variation of weather for identification and attribute estimated effects to alternative channels. In doing so, we address additional gaps in existing literature by more accurately identifying the exposure of interest with high resolution data, evaluating more comprehensively potentially relevant weather variables, by accounting for covariates that vary across time and region, and by studying lag times and non-linear associations (Romero Starke *et al* 2021).

The remainder of the paper is structured as follows. In section 2, we present the data and methods. In section 3, we show and interpret our results. We conclude with a discussion and conclusions in section 4.

2. Methods

2.1. Overview

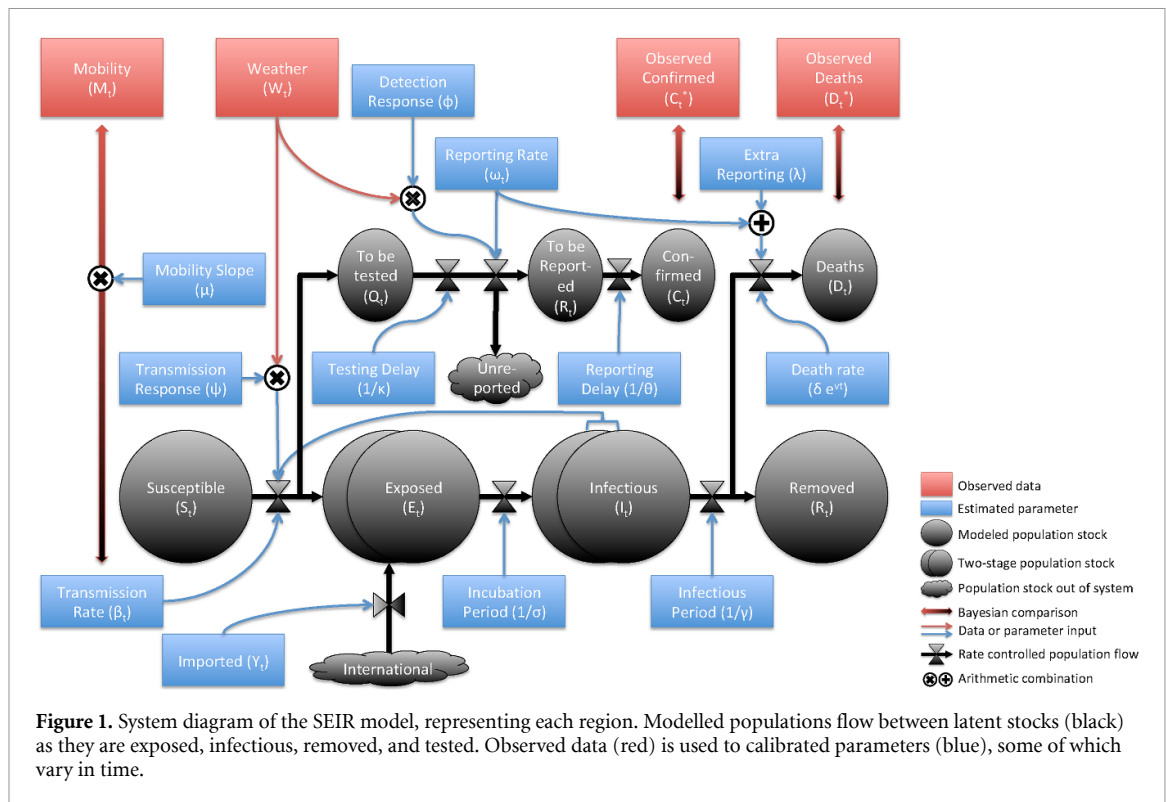
To shed new light on the effect of weather on COVID-19, we integrate two complementary approaches and leverage their individual strengths: econometric analysis with a panel regression model to empirically identify effects of weather on the disease and SEIR modelling to distinguish otherwise conflated epidemiological dynamics.

First, we use a panel regression model with distributed lags to identify effects of weather on COVID-19 dynamics. Similar to econometric studies on the effects of climate change that use random fluctuations of weather from year to year to learn about the effect of gradual changes of temperature over several decades (Hsiang 2016), our models uses random fluctuations of weather from day to day to learn about the effect of gradual changes of weather between seasons. To do so, our model includes a variety of fixed effects that absorb possibly confounding variation in the data, such as slow-moving trends, seasonality, and date-specific events.

To properly represent the epidemiological dynamic within a regression setting, we translate reported cases into a measure of instantaneous growth of the infectious population, derived from SIR dynamics, which has the same long-term growth rate as confirmed cases. We also use this approach to understand the effect of weather on mortality and to examine the role of mobility as a possible channel of the effect of weather on COVID-19 dynamics.

The panel regression allows us to identify key weather variables that influence transmission, navigating concerns of multicollinearity and omitted variable biases. The choice of influential weather variables and multicollinearity are in detail discussed in the SI. Regarding causality and other possible biases, weather can be considered exogenous to the evolution of the disease, ruling out reverse causality. Furthermore, the inclusion of a variety of fixed effects, including unit fixed effects and country-by-day fixed effects, rules out that time-constant differences between units and variables that change from day to day at the country level confound the estimated relationships. We also include unit-by-week fixed effects. These fixed effects will absorb, for example, variation in epidemiological dynamics due to health policies at the country or subnational level.

Second, we incorporate the weather variables identified by the regression model into an SEIR model, based on Kucharski *et al* (2020) (see SI figure 1). We also use insights obtained from the regression analysis on the role of mobility, the effect of weather on mortality, and the roughly linear association between weather and the growth of infectious population to inform how weather enters the SEIR model. The SEIR model uses both reported cases and deaths to simultaneously model an effect of weather on transmission and detection of new cases and thus to disentangle these two channels. The model also allows us to explore regional heterogeneity. The model is fit separately to each region, and a global average effect is obtained with a Bayesian hierarchical meta-analysis.



2.2. Data sources

We apply these methods to a high-resolution dataset with 4564 national and subnational units. These subnational units correspond to the first and second administrative levels in most countries. We construct this dataset based on several international and national epidemiological datasets, ERA5 reanalysis (Hoffmann *et al* 2019), and Google's Community Mobility report (Google 2020). Our dataset covers the year 2020, before the beginning of vaccinations and the dominance of new variants of the virus.

Our main source of epidemiological data is the repository maintained by John Hopkins University, which includes national counts for most countries and subnational counts for Australia, Canada, China, and the USA. We complement these data with several other sources of subnational counts of confirmed infections and deaths. In total we use subnational data for 33 countries and national data for 130 countries, which provides us with 4564 spatial units of observation (see SI section 1 for visualisations of the geographical and temporal coverage of the dataset).

Daily mobility data is obtained from Google's Community Mobility Report (Google 2020). We conduct a principal component analysis to reduce dimensionality and use the first principal component of the mobility categories unless stated otherwise. Data on governance indicators come from the GovData360 compendium which contains 33 datasets with worldwide coverage and covering time spans of more than 10 years (WBG 2020).

Weather data is obtained as ERA5 reanalysis provided by the European Centre for Medium-Range Weather Forecasts (Hoffmann *et al* 2019). The data consists of weather variables on a regular grid with a horizontal resolution of 0.25° latitude by 0.25° longitude.

2.3. Synthetic infectious population

In our dataset confirmed infections of COVID-19 are recorded as cumulative confirmed cases C_t reported up to a specific date t . For the panel regression we use the disease growth rate as the dependent variable, but the growth rate in confirmed cases loses statistically useful variation after a region's first wave of infections. Instead, we use the SIR model as a theoretical framework (SI section 2) to calculate the synthetic infectious population J_t as:

$$J_{t+1} = C_{t+1} - C_t + (1 - \gamma)J_t,$$

with a constant recovery rate γ . We take the value of γ from Kucharski *et al* (2020) as 1/8.1 days. The J variable is an estimate of the currently infected population (cumulative confirmed cases minus the recovered population). Importantly, in the case of continued growth, the value of γ does not affect the estimated growth rate, and the estimated growth rate equals the growth rate of confirmed cases.

2.4. Panel regression model

We estimate a panel regression model with several lags of weather and mobility data. Because weather and mobility are strongly serially correlated, we use averages over three consecutive days. That is, we predict the growth rate of infectious population on day t with average weather and mobility over the time period $[t-1, t-3]$, $[t-4, t-6]$, ... In our main specification, the dependent variable is the growth rate of synthetic infectious population calculated as the first-difference of the logarithm of the number of synthetic cases. Furthermore, we include a set of indicator variables to control for possible unobserved confounders. The full model can be written as:

$$y_{it} = \sum_{j=1}^4 \sum_{l=1}^L X_{i,j,t-l} \beta_{j,l} + \nu_{i,d} + \xi_{s(i),t} + \gamma_{i,w} + \epsilon_{it}, \quad (1)$$

with $y_{it} = \log(J_{i,t}) - \log(J_{i,t-1})$ and $J_{i,t}$ the total number of synthetic infectious population of unit i at date t , a vector of coefficients β_l and a matrix of observations $X_{i,j,t}$ (including lagged averages of weather and mobility) for every weather variable indexed by j , and unit-by-day-of-week fixed effects $\nu_{i,d}$, date-by-spatial-superset⁴ fixed effects $\xi_{s(i),t}$ and unit-by-week fixed effects $\gamma_{i,w}$. We include these fixed effects to absorb any time-invariant variation between locations and to flexibly account for trends over time. Equation (1) implicitly includes intercepts. For example, the unit-by-day-of-week fixed effects can be interpreted as implicitly including unit-specific intercepts. The instantaneous effect is excluded (i.e. time lag $l=0$) because we assume that in 2020 it took at least one day (or night) for a detected case to be reported and hence to appear in our reported cases. Errors ϵ_{it} are clustered by the unit of observation to account for heteroskedasticity and serial correlation.

In order to reduce the non-stationarity of our dependent and explanatory variable, we estimate a first-differenced version of this model, which also reduces multicollinearity below critical thresholds (SI section 3.1).

We also use the estimated coefficients of the panel regression model to illustrate the effect of a persistent change of temperature ΔT on the number of daily confirmed cases. To do so, we assume a constant infectious population $J_t = J_0 = 1$ in the counterfactual scenario and set all coefficients except the coefficient for temperature to zero:

$$J_{t+1} = J_t \exp \left(\sum_{l=0}^L \beta_l \Delta T \right). \quad (2)$$

We then calculate the number of confirmed cases as:

$$C_{t+1} = C_t + J_{t+1} - J_t (1 - \gamma), \quad (3)$$

using the definition of synthetic infectious population introduced above.

2.5. SEIR model

The SEIR model divides the population of each spatial unit into a susceptible population (S_t), stages 1 and 2 exposed populations (E_{1t} and E_{2t}), stages 1 and 2 infectious populations I_{1t} and I_{2t} , and a removed population. In this context, the attribute exposed describes an individual that has been infected by another individual, but is not yet infectious. Initially, $S_{t=0} = N$, the population of the spatial unit. The newly exposed cases (entering stage 1 of exposure) on day t is calculated as:

$$X_t = \exp \left(\log \tilde{\beta}_{t-1} + \chi_{1,\text{mod } 7} + \psi \cdot w_{t-1} \right) S_t (I_{1,t-1} + I_{2,t-1}) / N,$$

where $\tilde{\beta}_t$ is a smoothly varying underlying transmission rate, $\chi_{1,d}$ is a day-of-week effect, and ψ is a vector of weather coefficients, and w_t is a vector of weather values demeaned by spatial unit and scaled to a unit variance globally. The underlying observation rate captures unexplained shifts in behaviour and policy not driven by weather. The transition rate of populations from exposed to infectious is σ , the inverse of the incubation period, and the transition rate from infectious to removed is γ , the inverse of the infectious period.

⁴ In the case of a subnational observation, e.g. a city or a region, the spatial superset is the corresponding country. In case of a country, the superset is the world.

As individuals move from stage 1 to stage 2 of exposure, they also become ‘reportable’, and transition out of the reportable classification at a rate κ , the inverse of the pre-testing period. The portion of reportable people who are tested is the observation rate,

$$\omega = \exp(\log \tilde{\omega}_{t-1} + \chi_{2, \text{tmod}} + \phi \cdot w_{t-1}),$$

where $\tilde{\omega}_t$ is a smoothly varying underlying observation rate, capturing unexplained shifts in under-reporting. The tested population is ultimately reported as confirmed cases at a rate of θ , the inverse of the reporting delay.

We fit the model using the computational Bayesian MCMC system Stan (Gelman *et al* 2015), in two stages. The model above is calibrated to each spatial unit independently, producing an initial estimate of the parameters. We then apply a Bayesian hierarchical model to meta-analyse these parameter values within each country.

3. Results

3.1. Global effects of weather on COVID spread

Exploratory analysis with our panel regression model suggests that the effect of weather can be estimated well using a model with four meteorological variables (SI section 3): air temperature, thermal comfort as measured by the universal Thermal Climate Index (UTCI, which accounts for temperature, humidity, and clothing choices adapted to weather), solar radiation, and precipitation. We also test models that include alternative measures of humidity, but find that collinearity between temperature and humidity prevents reliable identification of their individual effects (SI section 3.2).

We find statistically significant effects of weather (figure 2) on the growth of infectious population. These effects of weather on disease growth are approximately linear in the level of the weather variables except for temperature (SI section 3.4). For temperature, disease growth is lowest at intermediate levels. We estimate models with 5 lags (15 days), 10 lags (30 days), and 20 lags (60 days). The estimated cumulative effects tend to increase in magnitude as we add lags to the model (see also SI section 3.3). Simulations with our SEIR model suggest that these long delays between a weather shock and its effect on the growth of confirmed cases can be explained with the dynamics around testing (see below). Indeed, we find very similar temporal structures of the response to a weather shock once we add testing to the model.

For temperature, the cumulative effect over the first 15 days increases slightly if we add additional lags (SI figure 13), which we attribute to the relative persistence of this variable. Without the additional lags, the results are more conservative. We hence focus on the effect over the first 15 days, as suggested also by previous work on the effect of weather on COVID-19 (Carleton *et al* 2020). Further below, we examine the temporal structure of the response to changes in weather in more detail using simulations with our SEIR model.

Precipitation has the largest and most statistically significant cumulative effect over 15 days. Cumulative over 15 days an increase of temperature, UTCI, solar radiation, and precipitation by one unit (K, index point, W m^{-2} , mm d^{-1}) changes the day-to-day growth rate of the infectious population by -0.24 [-0.0075 , 0.0027] (point estimate and 95% CI), -0.13 [-0.0032 , 0.0006], $+0.01$ [-0.0002 , 0.0005], and -0.4 [-0.0063 , -0.0017] percentage points, respectively.

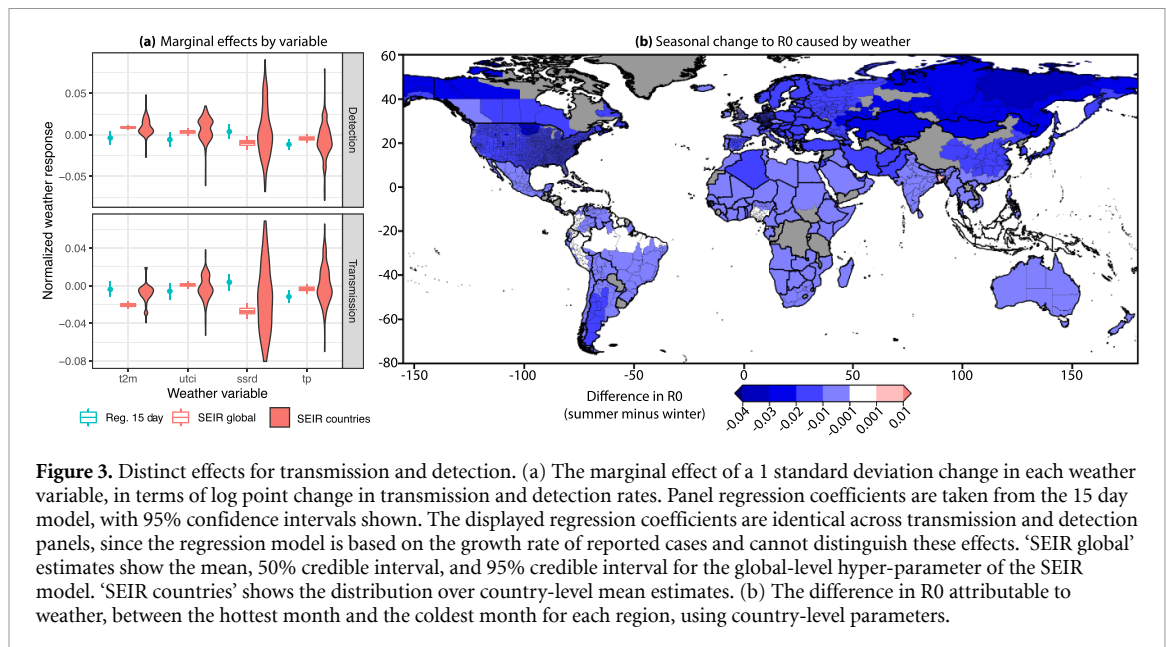
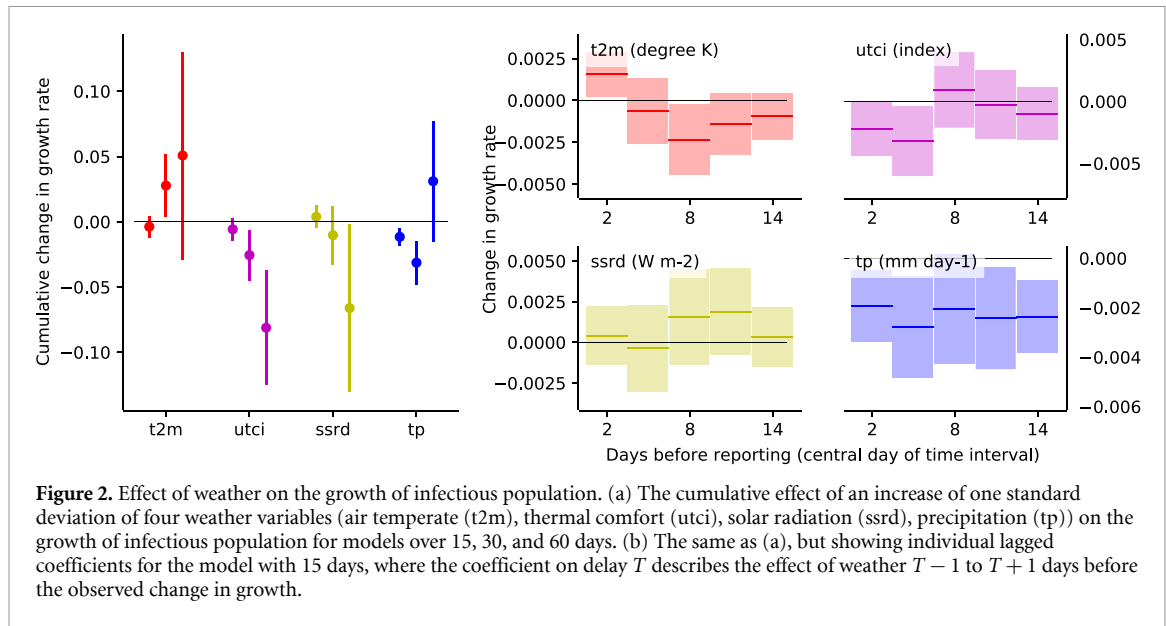
These effects may appear small, but they accumulate over time if weather changes persistent. This can result in large changes of the number of confirmed cases. For example, if temperature is lower by 5°C over 90 days relative to a counterfactual scenario, at the end of the period the number of daily confirmed cases is almost twice as large (190%) as the number in the counterfactual.

3.2. Distinct transmission and detection channels

We find that weather on a given day has a significant effect on disease growth on the immediately following 3 days, as well as growth rates up to at least 2 months into the future. The existence of both immediate and long-lasting effects in figure 2 has implications for the transmission and detection channels. If weather only affected transmission, its effect would not emerge until the second lag (days 4–6), while conversely effects on detection should be more immediate.

To study these channels further, we incorporate the above four weather variables into an SEIR model. In the calibrated model, detection effects are found to occur within 3 days of a weather shock and show a rebound effect as additional (reduced) detections reduce (increase) the pool for later detection. Transmission effects last longer, due to feedback dynamics and have no rebound effect (see SI figure 21).

Temperature, UTCI, solar radiation, and precipitation are all estimated to have statistically significant but small effects on detection (figure 3(a)). A shift in these three weather variables by one unit (K, index point, W m^{-2} , mm d^{-1}) results in an instantaneous change in the detection rate of 0.068 [0.031 – 0.11]%, 0.029

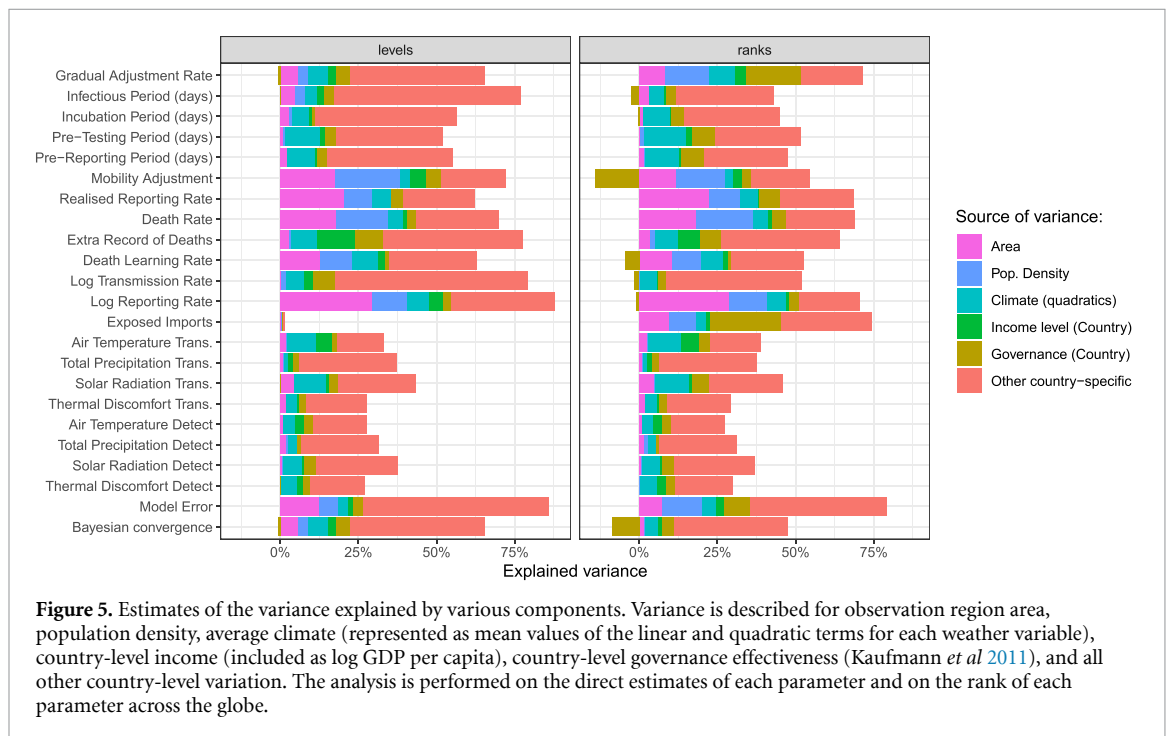
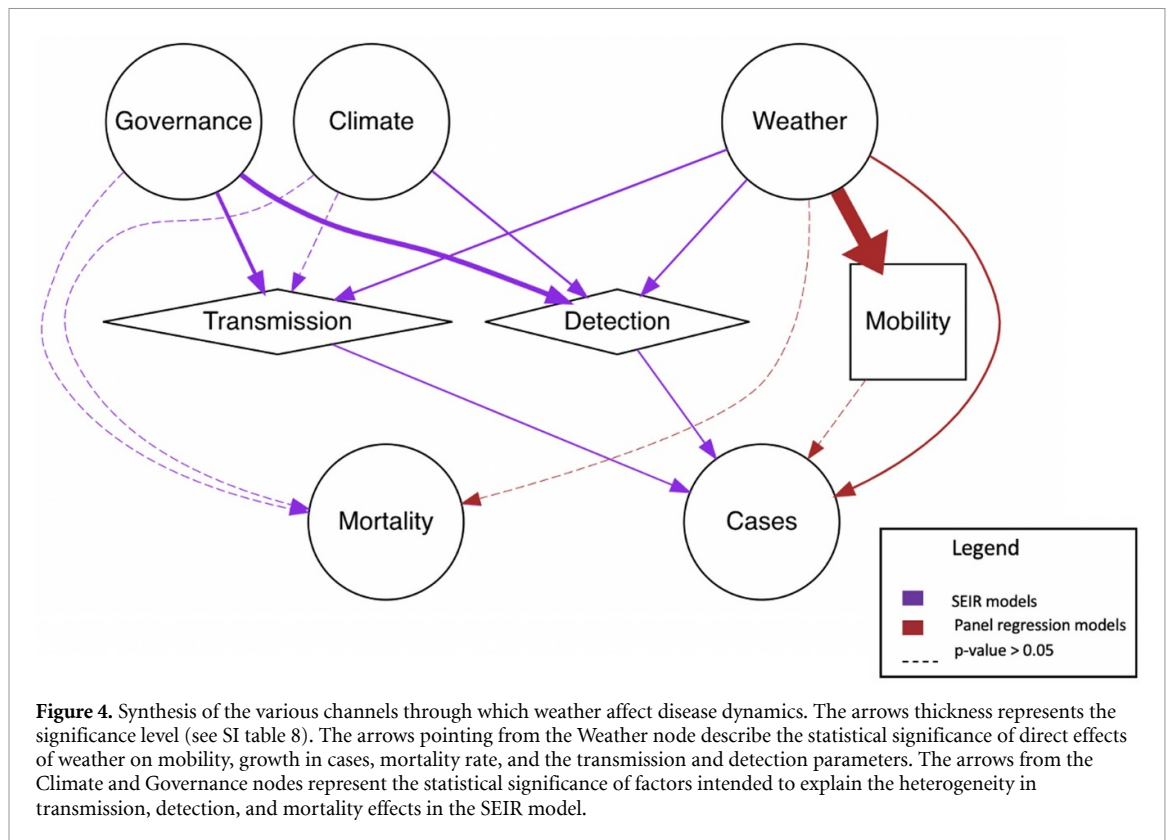


[0.0071–0.051]%, $-0.011 [-0.02--0.0025]\%$, and $-0.084 [-0.16--0.00098]\%$, respectively. Temperature and solar radiation also affect transmission, with an increase by one unit resulting in an instantaneous change in the transmission rate of $-0.22 [-0.27--0.16]\%$ and $-0.025 [-0.033--0.018]\%$. While these values are not equivalent to the estimates of the panel regression, as they reflect changes in underlying rate parameters, the signs of these effects are consistent with the 15 day panel results.

There is considerable variation in the estimated effects across countries. The interquartile range across countries of the estimated effect of a 1 K change in daily temperature on transmission rates is -0.039% to 0.0055% . Figure 3(b) combines this variation with weather variation to describe total seasonal changes in R_0 . To summarize the magnitude of the effect of weather on local epidemiological dynamics, we calculate the variation in transmission rates over time attributable only to weather. On average, the standard deviation of weather-driven variation in transmission is 4.9% of the mean transmission rate, with a 95% interval across countries from 1.7% to 9.6% (SI figure 22).

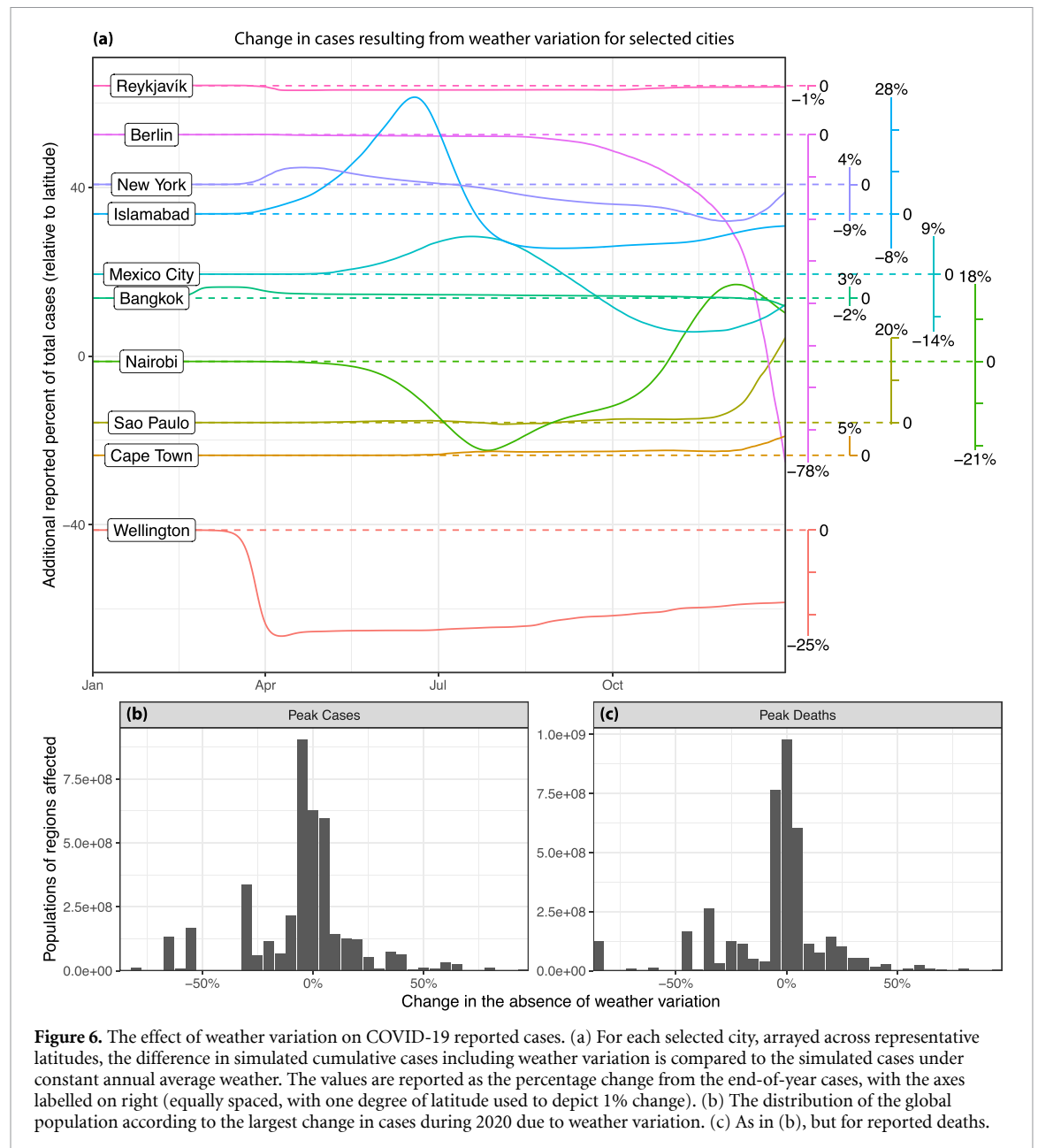
3.2.1. Mechanisms connecting weather with disease transmission

We use our estimated coefficients of the panel regression model and the SEIR model and conduct statistical tests to compare the importance of individual channels (figure 4). By comparing model skill with and without weather, we find significant effects of weather on both transmission and detection. These effects do



not reflect changes in mobility: while we do find statistically significant support for an effect of weather on mobility, we do not find corresponding evidence for an effect of mobility on reported cases.

The SEIR model allows us to explore the geographical heterogeneity of the different channels. Specifically, we focus on how climate and governance moderate the effect of weather on epidemiological dynamics (see figure 5). The heterogeneity in weather effects can be partly explained by adaptation to annual climate (up to 10% of the variation in weather effects), income levels and governance (up to 5% each). The remaining variation is idiosyncratic to each country or region. Climate is a strong explanatory factor for several other model parameters, with the strongest effect on reporting delays (pre-testing period and



pre-reporting period), explaining 15%–20% of their global heterogeneity, although the majority of the variance in these parameters is unexplained by climate, governance, or income.

We find that climate significantly moderates the effect of weather on detection ($p < 0.05$) but not on transmission. This could point to differences between the effect of weather on transmission and detection, with detection benefitting more from being modelled as a non-linear function of weather than transmission. Regarding governance, we find highly significant moderation of both detection and transmission ($p < 0.001$).

3.2.2. Improvements in the predictiveness of disease dynamics

While the instantaneous effects of weather on transmission are small, the epidemiological process includes a strong positive feedback loop which compounds them. As shown in figure 6 for cities selected across a range of latitudes, weather can result in both increases and decreases of infections relative to a counter-factual in which weather is kept constant. These effects occur in both tropical and temperate and in both rich and poor countries. Some cities show a seasonal cycle (New York, Islamabad, Mexico City, Nairobi), while others show little differences between the modelled dynamic and the counterfactual (Reykjavik, Bangkok, Cape Town). The largest change is shown in Berlin, where cases are estimated to be 78% less due to weather. This kind of ‘run-away’ dynamic could reflect dynamic policy responses, which occur out-of-sync in the counterfactual

simulation. The total effect of weather on confirmed cases is further complicated by the timing of chance co-occurrence between weather changes and pandemic waves and lock-downs. On average, weather variation is projected to have adjusted peak case numbers by 10%, while 10% of the global population experienced increases greater than 20%.

The inclusion of weather data can also improve estimates of key epidemiological parameters, as models that assume no weather variation misattribute this variation to other sources. We calibrate a version of our SEIR model without weather variables and compare the estimates to our main model (SI 4.6). We find that the detection rate of our main model is about 12% lower than the estimated parameter in the model without weather, with the underlying reporting parameter 18 log points lower. This implies that considerably more people may have been infected with COVID-19 than prior estimates suggest. The gradual adjustment rate, which allows the underlying transmission rate to adjust over time to account for changes in policy and behavior, is also considerably lower, reflecting the tendency to misattribute the variation due to weather to policy changes.

4. Discussion

As the COVID-19 pandemic evolves from rapid outbreaks to cycles potentially adjusted by seasonality (Charters and Heitman 2021), there is a need for comprehensive models to reliably inform health policy with scenario analysis and predictions. This endemic scenario indicates that the future evolution of the virus will depend on many unpredictable factors such as the emergence of new variants (Phillips 2021, Sonabend *et al* 2021). In contrast, its seasonal cycle due to weather variation may be comparatively predictable (Paraskevis *et al* 2021). As susceptibility to the disease falls, the role of environmental factors is expected to increase, in a way similar to influenza (Carlson *et al* 2020, Moriyama *et al* 2020, Kronfeld-Schor *et al* 2021).

In our study, random variation in weather provides a natural experiment to identify the effect of weather on infectious population and confirmed cases. The empirical results of our panel regression model show that weather has statistically and epidemiologically significant effects on the spread of COVID. While the instantaneous effect of a day with unusual weather is small, our results suggest that seasonal shifts can be important drivers of the number of infected individuals. We then use an SEIR model to isolate major channels of these effects. Our results suggest that weather influences both transmission and detection, which has important implication of the interpretation of the results of previous observational studies (e.g. Menebo 2020, Runkle *et al* 2020, Tosepu *et al* 2020) as the response of detection to weather can confound the observed effect of weather on reported cases.

We find that higher temperatures and levels of solar radiation limit the spread of COVID-19. This association with temperature conforms to other recent results (Ganslmeier *et al* 2021), although considerable disagreement exists on the temporal dynamics, non-linearity, and universality of weather effects (Yao *et al* 2020, Islam *et al* 2021, Sera *et al* 2021). We show that the effects of a weather shock continue to build over months, suggesting that current systems do not react to counter-balance these effects. Modest variations in R_0 can lead to considerable shifts in cases, deaths, and strains on hospital facilities.

Attributing changes in the dynamic of the disease to fluctuations or gradual changes in weather is a challenge for statistical analysis. Our results suggest that statistical attribution is difficult also because of long time delays between a weather shock and the effect on the growth rate of confirmed cases. With simulations with the SEIR model we can attribute such delays, which our analysis suggests can be up to several weeks long, to the dynamics around testing and reporting of cases.

The COVID-19 pandemic has been influenced by weather variation across the globe. We also find that there is considerable heterogeneity in the effects of weather, driven both by differences in the strength of the seasonal cycle and the underlying sensitivity of transmission to weather. Furthermore, regional differences in transmission sensitivity are partly explained by strength of governance and climate adaptation, but much of the variation is idiosyncratic to each country. Our results therefore also suggest that global average estimates of both the effect of weather and other epidemiological parameters can be misleading, highlighting the importance of targeted, contextual policies.

Our results can be used to improve the timing of targeted interventions. A weather shock today will have observable effects for 1 (temperature, rainfall) to 2 (thermal comfort, solar radiation) months. Accounting for the effects of weather variation on detection and transmission can help explain shifts in R_0 and inform expectations about future case loads.

Our results come with some important caveats. First, the data on COVID cases and deaths is incomplete and suffers from several limitations. Most notably, our data has subnational coverage for only a few countries and not for all countries we have data for all months of 2020, which limits the representativeness of our

results for specific geographies and time periods. Furthermore, our results for countries with only national data should be interpreted with caution. Second, we have not included the effect of vaccinations or variants in our model since our data ends with the 2020 calendar year. The combined effect of new variants, new policies, and new immune system reactions since 2020 could significantly affect both the magnitude of weather effects and the contribution of different weather variables. Third, we do not account for different demographic groups, occupations, or other risk factors. Fourth, we are describing the effect of weather on reported cases, which include both false positive results and considerable missing data. Our estimate of the effect of weather on detection encompasses all effects that weather has on the reliability of COVID data (Cohen *et al* 2020).

Data availability statement

All code for reproduction is available at <https://github.com/openmodels/coronaclimate>. Input weather and collated records data, and estimated model outputs are available at <https://doi.org/10.5281/zenodo.7262562>.

Acknowledgment

We would like to thank Yang Liu, Susana Mourato, Sefi Roth, and participants of the session COVID-19: Earth Observations and Environmental Modeling Applications for Understanding, Quantifying, and Predicting the Seasonality (or Not) and Propagation of COVID-19 at the AGU 2020 fall meeting. M.L. acknowledges financial support by the UK's Economic and Social Research Council (ESRC) with grant number 2300776. A.D.M. acknowledges financial support by the PhD Studentship programme of the London School of Economics (LSE).

ORCID iDs

James Rising  <https://orcid.org/0000-0001-8514-4748>

Manuel Linsenmeier  <https://orcid.org/0000-0002-7790-4011>

Ana De Menezes  <https://orcid.org/0000-0001-8456-2470>

References

- Bukhari Q, Massaro J M, D'agostino R B and Khan S 2020 *Int. J. Environ. Res. Public Health* **17** 5399
- Carleton T, Cornet J, Huybers P, Meng K C and Proctor J 2020 *Proc. Natl Acad. Sci.* **118** e2012370118
- Carlson C J, Gomez A C, Bansal S and Ryan S J 2020 *Nat. Commun.* **11** 1–4
- Chan K H, Peiris J M, Lam S, Poon L, Yuen K and Seto W H 2011 *Adv. virol.* **2011** 1–7
- Charters E and Heitman K 2021 *Centaurus* **63** 210–24
- Chin A, Chu J, Perera M, Hui K, Yen H L, Chan M, Peiris M and Poon L 2020 Stability of SARS-CoV-2 in different environmental conditions *Lancet Microbe* **1** e10
- Cohen F, Schwarz M, Li S, Lu Y and Jani A 2020 The challenge of using epidemiological case count data: the example of confirmed COVID-19 cases and the weather *Environmental and Resource Economics: Perspectives on the Economics of the Environment in the Shadow of Coronavirus* vol 76, pp 447–517
- Cools M, Moons E, Creemers L and Wets G 2010 *Transp. Res. Rec.* **2157** 22–28
- Ganegoda N C, Wijaya K P, Amadi M, Erandi K H and Aldila D 2021 *Sci. Rep.* **11** 1–16
- Ganslmeier M, Furceri D and Ostry J D 2021 *Sci. Rep.* **11** 1–7
- Gelman A, Lee D and Guo J 2015 *J. Educ. Behav. Stat.* **40** 530–43
- Google 2020 (available at: www.google.com/covid19/mobility/)
- Hoffmann L *et al* 2019 *Atmos. Chem. Phys.* **19** 3097–124
- Hsiang S 2016 *Annu. Rev. Resour. Econ.* **8** 43–75
- Huang Z, Huang J, Gu Q, Du P, Liang H and Dong Q 2020 *Sci. Total Environ.* **736** 139487
- Islam N, Bukhari Q, Jameel Y, Shabnam S, Erzurumluoglu A M, Siddique M A, Massaro J M and D'Agostino Sr R B 2021 *Environ. Res.* **193** 110355
- Kaufmann D, Kraay A and Mastruzzi M 2011 *Hague J. Rule Law* **3** 220–46
- Kronfeld-Schor N, Stevenson T, Nickbakhsh S, Schernhammer E, Dopico X, Dayan T, Martinez M and Helm B 2021 *J. Biol. Rhythms* **36** 35–54
- Kucharski A J *et al* 2020 *Lancet Infect. Dis.* **20** 553–8
- Landier J, Paireau J, Rebaudet S, Legendre E, Lehot L, Fontanet A, Cauchemez S and Gaudart J 2021 *Sci. Rep.* **11** 1–13
- Leclercq I, Batejat C, Burguière A M and Manuguerra J C 2014 *Influenza Other Respir. viruses* **8** 585–6
- Majumder P and Ray P P 2021 *Sci. Rep.* **11** 1–10
- Mecenas P, Bastos R T d R M, Vallinoto A C R and Normando D 2020 *PLoS One* **15** e0238339
- Menebo M M 2020 *Sci. Total Environ.* **737** 139659
- Meyer A, Sadler R, Faverjon C, Cameron A R and Bannister-Tyrrell M 2020 *Front. Public Health* **8** 367
- Moriyama M, Hugentobler W J and Iwasaki A 2020 *Annu. Rev. Virol.* **7** 83–101
- Nouvellet P, Bhatia S, Cori A, Ainslie K E, Baguelin M, Bhatt S, Boonyasiri A, Brazeau N F, Cattarino L and Cooper L V *et al* 2021 *Nat. Commun.* **12** 1–9
- Paraskevis D, Kostaki E G, Alygizakis N, Thomaidis N S, Cartalis C, Tsiodras S and Dimopoulos M A 2021 *Sci. Total Environ.* **768** 144578

- Phillips N 2021 *Nature* **590** 382–4
- Romero Starke K, Mauer R, Karskens E, Pretzsch A, Reissig D, Nienhaus A, Seidler A L and Seidler A 2021 *Int. J. Environ. Res. Public Health* **18** 6665
- Runkle J D, Sugg M M, Leeper R D, Rao Y, Matthews J L and Rennie J J 2020 *Sci. Total Environ.* **740** 140093
- Sera F *et al* 2021 *Nat. Commun.* **12** 1–11
- Seyer A and Sanlidag T 2020 *Lancet Microbe* **1** e8–e9
- Shao W, Xie J and Zhu Y 2021 *Environ. Res.* **194** 110608
- Sonabend R *et al* 2021 *Lancet* **398** 1825–35
- Tosepu R, Gunawan J, Effendy D S, Ahmad L O A I, Lestari H, Bahar H and Asfian P 2020 *Sci. Total Environ.* **725** 138436
- Walrand S 2021 *Sci. Rep.* **11** 1–9
- Wang D, Wu X, Li C, Han J and Yin J 2022 *Sci. Total Environ.* **826** 154182
- WBG 2020 (available at: <https://govdata360.worldbank.org/>)
- Wilson D J 2021 *Weather, Mobility and COVID-19: A Panel Local Projections Estimator for Understanding and Forecasting Infectious Disease Spread* (Federal Reserve Bank of San Francisco)
- Wu Y, Jing W, Liu J, Ma Q, Yuan J, Wang Y, Du M and Liu M 2020 *Sci. Total Environ.* **729** 139051
- Xu R, Rahmandad H, Gupta M, DiGennaro C, Ghaffarzadegan N, Amini H and Jalali M S 2021 *Lancet Planetary Health* **5** e671–80
- Yao Y, Pan J, Liu Z, Meng X, Wang W, Kan H and Wang W 2020 *Eur. Respir. J.* **55** 2000517
- Yechezkel M, Weiss A, Rejwan I, Shahmoon E, Ben-Gal S and Yamin D 2021 *BMC Public Health* **21** 1–13