

Worldwide Correlations Support COVID-19 Seasonal Behavior and Impact of Global Change

Nicolas Hernandez and Gustavo Caetano-Anollés

Department of Crop Sciences and Carl R. Woese Institute for Genomic Biology, University of Illinois, Urbana, IL, USA.

Evolutionary Bioinformatics
Volume 19: 1–12
© The Author(s) 2023
Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/11769343231169377



ABSTRACT: Many viral diseases exhibit seasonal behavior and can be affected by environmental stressors. Using time-series correlation charts extrapolated from worldwide data, we provide strong support for the seasonal development of COVID-19 regardless of the immunity of the population, behavioral changes, and the periodic appearance of new variants with higher rates of infectivity and transmissibility. Statistically significant latitudinal gradients were also observed with indicators of global change. Using the Environmental Protection Index (EPI) and State of Global Air (SoGA) metrics, a bilateral analysis of environmental health and ecosystem vitality effects showed associations with COVID-19 transmission. Air quality, pollution emissions, and other indicators showed strong correlations with COVID-19 incidence and mortality. Remarkably, EPI category and performance indicators also correlated with latitude, suggesting cultural and psychological diversity in human populations not only impact wealth and happiness but also planetary health at latitudinal level. Looking forward, we conclude there will be a need to disentangle the seasonal and global change effects of COVID-19 noting that countries that go against the health of the planet affect health in general.

KEYWORDS: Active cases, correlation analysis, Environmental Performance Index, latitude, mortality, SARS-CoV-2, seasonal forcing, State of Global Air, time series

RECEIVED: December 26, 2022. **ACCEPTED:** March 14, 2023.

TYPE: Original Research

FUNDING: The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: Office of Research and Office of International Programs, College of Agricultural, Consumer and Environmental Sciences at the University of Illinois at Urbana-Champaign and New Frontiers Initiative, National Center for Supercomputing Applications (NCSA).

DECLARATION OF CONFLICTING INTERESTS: The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

CORRESPONDING AUTHOR: Gustavo Caetano-Anollés, Department of Crop Sciences, University of Illinois at Urbana-Champaign, 332 NSRC, 1101W Peabody Drive, Urbana, IL 61801, USA. Email: gca@illinois.edu

Introduction

The Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2) continues to spread since the COVID-19 disease was first reported on December 2019 in Wuhan, China.¹ As of February 15, 2023, more than 670 million cases and 6.8 million deaths have been reported worldwide.² Its unrelenting spread has been accompanied by the appearance of numerous viral variants with increasingly complex constellations of mutations, some of which are jeopardizing vaccination and testing performance.³ As with SARS-CoV, MERS-CoV and Influenza A viruses, SARS-CoV-2 spreads primarily via respiratory droplet and aerosol transmission, which occurs in close-contact situations where a person is talking, coughing, or sneezing.^{4,5} Prolonged exposure to an infected person or brief exposure to a symptomatic individual is associated with a higher risk of transmission.^{4,5} Coming into contact with a contaminated surface is another proven transmission mechanism, as SARS-CoV-2 has been shown to survive up to 28 days on certain surfaces.⁶ Symptom onset occurs an average of 5 days after contracting the virus, but in some cases it can be as long as 14 days.⁵ While the most common symptoms include fever, headaches, cough, anosmia and ageusia,⁵ a wide range of clinical manifestations of the disease have been recently dissected with machine learning approaches.⁷ Note that asymptomatic and pre-symptomatic individuals are still contagious and able to transmit the virus to others.⁸

The incidence of COVID-19 infections is influenced by the environment, expectedly though seasonal cycles and global change effects. Seasonal cycles of viral respiratory infections are

common.⁹ For instance, Influenza cases in the United States spike in the fall and winter, increasing in October, peaking in February, and declining in May. Chicken pox incidence similarly increases between winter and spring months and dies down by summer. These trends have been mapped for over a century.^{10,11} In contrast, COVID-19 is a novel infectious disease. Consequently, a disease-associated seasonal cycle has not been thoroughly explored. A combination of factors are responsible for seasonal epidemics, including direct environmental effects on viral transmissibility, initial susceptibility of the human population, and effects on immune response.⁹ During early stages, stable oscillations typical of seasonal patterns of disease cannot be accurately modeled, especially because environmental drivers cannot stop transmission during the rampant phase of a pandemic at a time when immune response is weak.¹² Despite limited knowledge, there is significant evidence supporting the seasonal behavior or SARS-CoV-2 infection.¹³ First, human coronaviruses are part of a group of RNA viruses colloquially known as “winter” viruses, 4 of which (NL63, 229E, OC43, and HKU1) have previously been identified as seasonal.¹⁴ Second, the half-life of the SARS-CoV-2 virus is affected by temperature and humidity, including on surfaces or under dry conditions.^{6,15,16} Moreover, smaller-scale studies of COVID-19 incidence also revealed a negative association between temperature and COVID-19 case occurrence. One study reported that in subtropical cities of Brazil a 1°C rise was associated with about 5% decrease in the number of daily cumulative cases when average ambient temperature was less than 25.8°C.¹⁷ Similarly, daily



new cases of COVID-19 reported in 166 countries decreased by 0.85% for every 1°C increase in temperature.¹⁸ Interestingly, multiple early pandemic case studies on COVID-19 showed conflicting results in correlation analysis. For instance, a study in Oslo, Norway reported positive correlations between higher temperatures and Covid-19 daily new cases,¹⁹ while another studying 122 cities in China found a 4.9% rise in daily confirmed cases for every 1°C increase in temperature when the mean temperature was less than 3°C.²⁰ Conflicting data can most likely be explained by early pandemic variability, which can be rectified by studying COVID-19 incidence for longer periods.²¹ Third, latitude, which is a strong predictor of temperature, was found to be correlated with COVID-19 epidemiological variables.^{13,22} It is well known that latitude affects the amount of solar radiation a location receives; on the equator the rays are most direct, yet the strength lessens as one moves toward either pole. This in turn affects the temperature of an area, as locations further from the equator receive less warming energy and are thereby colder on average. When studying a virus expected to thrive in colder temperatures, latitude can play a critical role in mapping case incidence.^{13,22} Finally, weak seasonal oscillation signatures of COVID-19 cases and deaths during the first year of the pandemic were extracted in sets of 5 countries in the Northern and Southern Hemispheres using ensemble mode decomposition (EEMD) methods.²³

Other drivers besides temperature and latitude may be of significance. We recently proposed that multiple seasonal drivers behind disease spread (and the spread of COVID-19 specifically) are in “trade-off” relationships and can be better described within a framework of a “triangle of viral persistence” modulated by the environment, physiology, and behavior.¹³ Trade-offs exist when one trait cannot increase without a decrease in another, especially when the environment affects physiologies and behavior through time and space-constrained limitations in matter-energy and information. In fact, a number of factors impinge on the multidimensional performance space of the triangle of viral persistence and could be relevant to COVID-19, including the effect of the following environmental health indicators:

(i) *Air Quality* has been shown to be a factor to consider in viral respiratory diseases.^{5,24,25} Correlation between *Air Quality* and community health outcomes has been discussed multiple times^{26,27} and is associated with a variety of quality-of-life indicators that may put less developed communities at a higher risk of contracting and dying from the disease. Air pollution in its various forms increases epithelial permeability to viral receptors, impairs phagocytosis and antigen presentation, and amplifies the virulence of respiratory infections, increasing host susceptibility to a range of respiratory viral diseases overall.²⁸

(ii) *Household Use of Solid Fuels* such as coal, kerosene, and biomasses (wood, charcoal, crop waste) is attributed to 3.8

million premature deaths each year,²⁹ and is a leading health risk factor in Asia, Africa, and Central/South America.³⁰ *Household Air Pollution* (HAP) is the encompassing term for byproducts of fuel burning that result in ambient pollutants. These ambient pollutants include CO, NO_x, and *Particulate Matter* (PM), which is comprised mainly of SO₂, NO₃, NH₄, organic and elemental Carbon, Silicon, and Sodium, at >1%.³¹ Concentrations of air pollutants can be exponentially higher in homes that use solid fuels compared to homes that use clean fuels and electricity, causing massive aggregations of health-impacting pollutants in small spaces over prolonged periods of time. Household use of solid fuels is linked to major respiratory diseases such as Acute Lower Respiratory Infections (Pneumonia), Chronic Obstructive Pulmonary Disease (COPD), and Chronic Lower Respiratory Disease.^{32,33} Exposure to HAP is responsible for 45% of all pneumonia deaths in children under 5 years of age and is a contributing factor to 28% of all adult deaths to pneumonia. A fourth of all adult deaths from COPD in low to middle-income countries are caused by prolonged exposure to pollution from usage of household solid fuels.²⁹ PM is another important metric for evaluating HAP and ambient outdoor pollution risk, measured as concentration of particulate matter under 2.5 μm (PM_{2.5}). Statistically significant links have been proved between the occurrence of acute respiratory distress syndrome (ARDS) and exposure to PM_{2.5} and other air pollutants.²⁷ Long-term exposure and inhalation of PM_{2.5} has been shown to be extremely detrimental to human health and was a contributing factor in 4.1 million deaths from heart disease, lung cancer, chronic lung disease, and respiratory infections worldwide in 2016.³⁴ In the European Union, the average life expectancy is estimated to be reduced by 8.6 months due to exposure and inhalation of PM_{2.5}.^{26,35} This is because PM_{2.5} and the chemicals it contains can induce free radical production and generate reactive oxygen species, as its surface is rich in iron, copper, zinc, manganese, and lipopolysaccharides, which decrease the antioxidant capacity of cells and cause oxidative stress on cells and DNA.²⁷ Concurrent peroxidation of lipids increases the amount of intracellular Calcium ion concentrations, which activate a series of inflammatory responses leading to cell damage, apoptosis, and necrosis. Inflammation and cell death can worsen certain comorbidities, such as asthma and pulmonary diseases.^{27,28}

(iii) *Ozone (O₃) concentration* is an important measure of air quality. O₃ is a major constituent of photochemical fog caused by reactions between NO₂, hydrocarbons, and UV light.²⁵ An increase in O₃ levels can cause oxidative stress in the respiratory tract similar to that of PM_{2.5}, triggering bronchial inflammation and hyper-responsiveness.²⁵ This in turn obstructs airway flow and pulmonary function. Short-term O₃ exposure is associated with an increased risk of COPD hospitalization and can exacerbate asthma

conditions.²⁵ In longitudinal studies, O₃ was significantly associated with emphysema and decreased lung function, while PM_{2.5} was not. This can be important. While there is regulation to prevent short-term rises in O₃ levels, long-term averages are not decreasing³⁶ and are not expected to decrease as climate change worsens.³⁷

In a previous study, epidemiological data collected worldwide during the first wave of the pandemic revealed weak but significant correlations between COVID-19 spread and both temperature and latitude, suggesting COVID-19 was seasonal.²² Our goal in this paper was to confirm these correlations with 18 more months of pandemic data, while mapping them month-by-month since the start of the pandemic. This effort is intended to provide a comprehensive view into how the strength of the overall correlation has changed over time. We also correlate epidemiological variables with the Yale's Environmental Performance Index (EPI),³⁸ which ranks countries against each other based on their performance toward certain environmental targets, thereby creating a global summary of the effect of "environmental health indicators" on COVID-19 incidence and mortality. Results are compared with data from the State of Global Air (SoGA) report.³⁴ Our analysis provides significant evidence supporting strong seasonal behavior and a likely role of global change parameters on the incidence of COVID-19.

Materials and Methods

Measurements of relationship

To assist in describing interactions between epidemiological variables and possible factors most accurately, Pearson product-moment correlation coefficients (r) were used to calculate variable relationships. This correlation is a measure of the strength and direction of a linear association between 2 variables, a covariance of both variables divided by the product of their standard deviations.³⁹ Spearman's rank-order correlation (r_s) was also used as a secondary measurement of relationship to account for the possibility of other monotonic relationships that Pearson correlation cannot capture as efficiently. This correlation is a measure of strength and direction of an association between 2 ranked variables. Both correlations result in values between -1 and 1 in which 1 is an exact positive correlation, -1 is an exact negative correlation. A value of zero implies no correlation between 2 variables. As such, if values approach zero the correlation is "weak," and if values approach |1| the correlation is deemed "strong."³⁹ The Breusch-Pagan (BP) heteroscedasticity test was used to determine if variance did not depend on auxiliary regressors in all correlations performed in this study. In other words, the BP test examines if residuals of linear regressions had changing variances. As expected in cross-sectional studies such as those of global change factors, we confirmed heteroscedasticity was common in correlations between the 32 EPI performance indicators and epidemiological

variables. This is because indicators embody a wide diversity of data with a large range of values for the different countries, including a large disparity between the largest and smaller values. To offset the problem we resorted to transformation of the dependent variable, acknowledging data manipulation complicates interpretation and selecting the transformation that minimized heteroscedasticity violations across EPI indicators following experimentation with log functions, square roots, and exponents. Correlation coefficients, associated p -values, and BP X^2 test statistic and corresponding p -values are provided in Supplemental Table S1.

COVID-19 epidemiological variables and definitions

All COVID-19 related data was retrieved from the *Our World in Data* COVID-19 Dataset,⁴⁰ which keeps daily records of cases, deaths, and vaccinations reported by each participating country. Data from 171 countries was retrieved, with 4 variables extrapolated from the dataset. To minimize frequency of repeating values (non-updated data), countries that did not update on a daily basis and/or had gaps in reporting were not included. To track epidemiological variables throughout the course of the pandemic, each variable was split on a month-by-month basis. Variables were as follows: (1) *Total cases per million at the end of each month (Total Cases)*, signifying the total population-weighted sum of cases reported in each country since the beginning of the pandemic to the last day of the specified month, (2) *Total deaths per million at the end of each month (Total Deaths)*, signifying the total population-weighted sum of deaths reported in each country since the beginning of the pandemic to the last day of the specified month, (3) *New cases per million at the end of each month (New Cases)*, signifying the population-weighted sum of reported cases between the first and last day of the specified month, and (4) *New deaths per million at the end of each month (New Deaths)*, signifying the population-weighted sum of reported deaths between the first and last day of the specified month. Values for 24 months (January 2020–December 2021) were retrieved for each variable, allowing construction of a time-series correlation plot for each variable represented above. This allowed to track the strength of each variable's relationship across a 2-year timespan.

Factor data variables and definitions

Factor data reside under 3 major categories: Temperature, Latitude, and Yale's 2020 EPI scores. Not all COVID-19 epidemiological variables were tested with each factor retrieved, as these base variables are used to test only for certain factors such that testing each base variable for all factors would not make statistical sense. The 4 categories are explained on a variable-to-variable basis:

Total Cases and Total Deaths variables described change in time of the total association between COVID-19 and a variety

of factors, including: (1) The **Average Yearly Temperature** of each country ($^{\circ}\text{C}$), (2) The **Latitude** of each country represented specifically by the geographic coordinates of its capital ($^{\circ}\text{N}$ or $^{\circ}\text{S}$ in absolute values), and (3) The **EPI score** of each country, which has values ranging 0 to 100. As the EPI does not cover the same range of countries as *Our World in Data*, we restricted our analysis to only to 151 countries for EPI-related data instead of the full 171. We also tested all EPI subcategories, which included Air Quality, Pollution Emissions, Climate Change, Environmental Health, and Sanitation scores for 1 month (March 2021) of the pandemic. This same month was also used to explore correlations with SoGA Report's raw Ozone, $\text{PM}_{2.5}$, and Household Solid Fuel Usage data. To summarize, Total Cases and Total Deaths were tested with average yearly temperature, latitude, and EPI scores. Additionally, Total Cases and Total Deaths data for the month of March 2021 was also tested with all EPI subcategories and SoGA data.

New Cases and New Deaths variables were only used to determine correlations with both temperature and latitude for each individual month of the pandemic. For temperature, we measured the relationship between the temperature of a specific month correlated only with the cases or deaths that were reported in that month. This allowed to determine peak periods when correlation is strongest while accounting for the difference in seasonal cycles between the northern and southern hemispheres. Unfortunately, average monthly temperatures of countries are difficult to compute, as most countries tend to have many geographic regions with varying average temperatures. As a solution, a city with retrievable temperature data was selected to represent our metric of **Average Monthly Temperature ($^{\circ}\text{C}$)**. To summarize; New Cases and New Deaths variables were tested with average monthly temperature and latitude. For example, "New Cases at the end of December 2020" takes the cumulative sum of all COVID-19 cases for each country between December 1st, 2020, and December 31st, 2020. Each resulting value is then divided by the population of its appropriate country and run through a correlation test alongside its selected city's average December temperature and latitude. The same is then done for every other month there is data for, with the resulting r value plotted on a graph. By doing this across 24 months, we uncovered global COVID-19 seasonal patterns showing the waxing and waning of the disease.

Results

A worldwide correlation analysis in the context of a time series of sequential observations offered a temporal view of correlations unfolding along the pandemic. We used both Pearson's r and Spearman's r_s coefficients. Pearson's r can be calculated without any assumptions, but inference on the strength of the association does require some assumptions be met: (1) that the data is derived randomly from, or is representative of, a population, (2) that both variables be continuous, jointly normally

distributed, and random, and that if there is a correlation, it is always linear, (3) that there are no relevant outliers, and (4) that every pair is measured separately from all other pairs.³⁹ There are various ways to deal with certain violations of these assumptions. Specifically, in the case of there being relevant outliers or a non-linear pattern, we can use Spearman's r_s , since it is more robust against outliers and does not require normally distributed data to make inferences on the strength of its association. Data can also be transformed to linearize a relationship, thereby allowing the data to meet Pearson's correlation's assumptions.

Our study showed increasingly stronger trends across all Total Cases and Total Deaths time-series correlation graphs for temperature and latitude. Total Cases approached correlation values of $r = -.450$ and $r_s = -.470$ for temperature and values of $r = .615$ and $r_s = .607$ for latitude. Total Deaths approached values of $r = -.350$ and $r_s = -.420$ for temperature and $r = .459$ and $r_s = .571$ for latitude. Selected bivariate Kernel density scatter plots illustrate the correlations and show countries located at latitudes spanning 30 to 60 $^{\circ}\text{N}$ or $^{\circ}\text{S}$ and average temperatures of 10 $^{\circ}\text{C}$ to 20 $^{\circ}\text{C}$ have been most impacted by the disease (Figure 1). The bimodal patterns in density plots suggest effects of planetary environmental factors. Remarkable cyclical patterns were uncovered in time-series correlation graphs, especially those of New Cases and New Deaths, that demand explanation. Most EPI performance indicators and categories showed weak to moderate negative correlation values under both cases and deaths, but these results could be confounded by an association between EPI indicators and latitude.

Main correlations

We found 3 types of seasonal and environmental correlations of significance impacting the incidence and severity of the COVID-19 pandemic:

- (i) *Correlations of Total Cases or Total Deaths with Temperature and Latitude.* Figure 2 shows the time series total correlation between Total Cases and Total Deaths at the end of each month and both average yearly temperature and latitude. In average yearly temperature correlations, there was a negative trend since the early months of the pandemic. The same was reflected in correlations with latitude, though at increasingly positive values. In other words, COVID-19 case incidence and deaths were globally correlated with lower temperatures and higher latitudes.

The correlation between Total Cases and average yearly temperature depicts a minor spike of statistically significant negative strength between February 2020 ($r = -.106$, $r_s = -.319$) and September 2020 ($r = -.053$, $r_s = -.266$), peaking in March 2020 at $r = -.278$, $r_s = -.441$ (Figure 2A). Only few countries exhibiting lower average temperatures had a higher number of population-weighted cases during the first few months of the

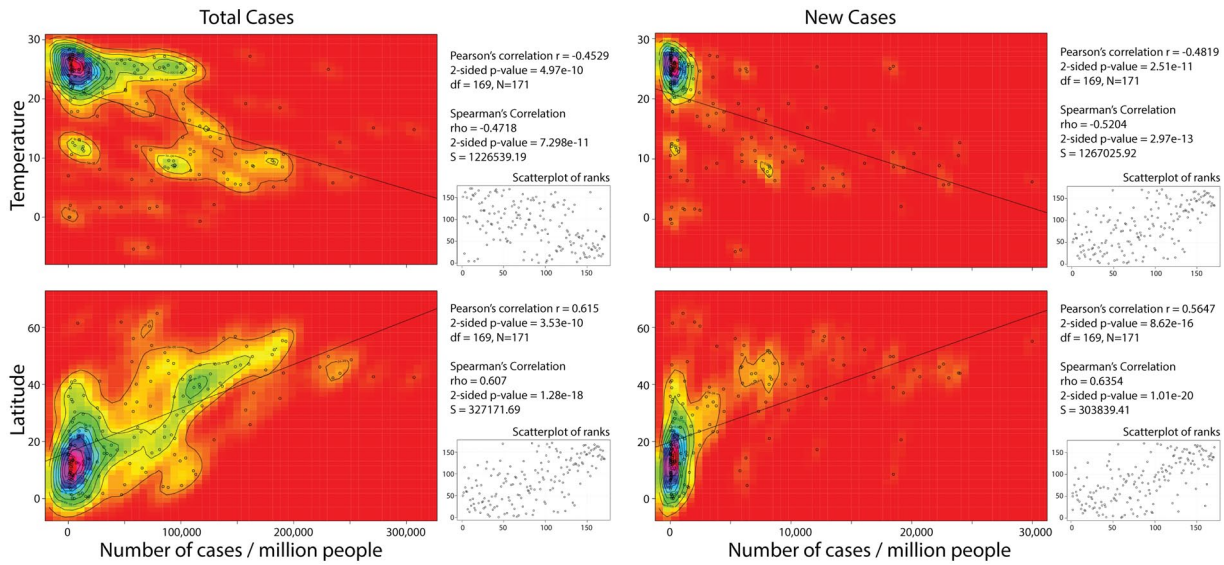


Figure 1. Example bivariate kernel density plots describe the relationship between the number of confirmed COVID-19 cases per million people and either temperature (°C) or latitude (°N or °S). Cases are given as Total Cases reported on December 2021 (left panels) and New Cases reported on December 2020 over each country's average yearly or monthly temperature, respectively, and over geographic coordinates of each country's latitude.

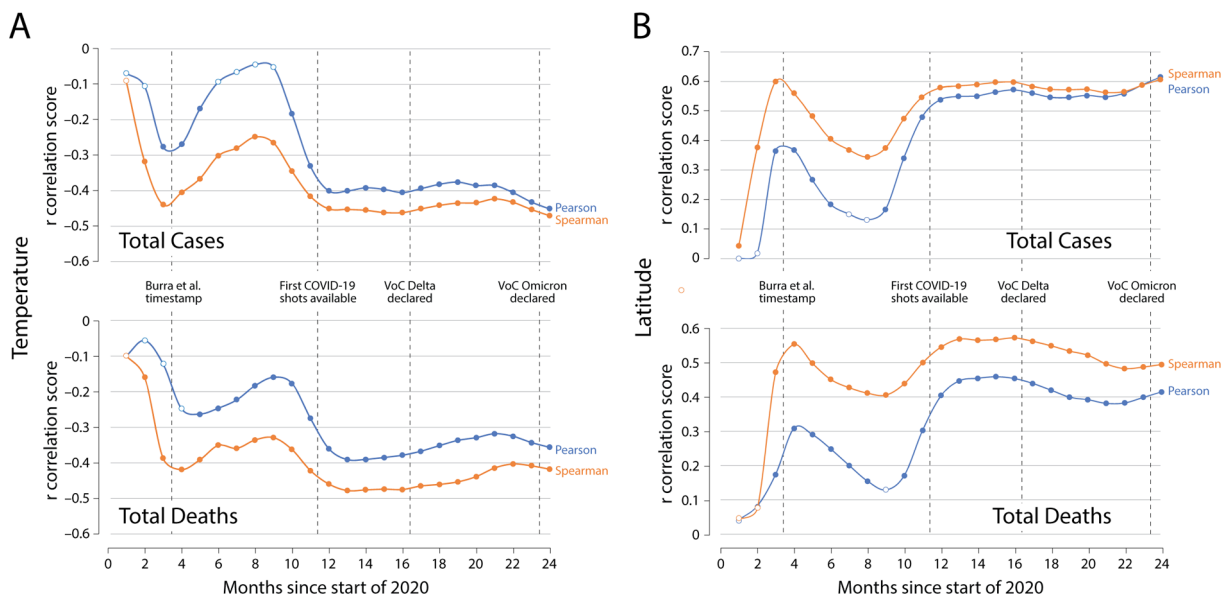


Figure 2. Time series correlation plot between Total Cases (top) and Total Deaths (bottom) and either average yearly temperatures (A) or latitude (B). Each point represents the r coefficient between the Total Cases or Total Deaths through each month and both yearly temperature averages or latitude. Closed symbols describe significant correlations with p -values less than .05. Important timestamps are indicated in the graphic including the work on seasonal behavior of Burra et al.²²

pandemic. By June 2020 almost all countries were affected by COVID-19, and the correlation was lost. Correlations between Total Cases and average temperature mostly even out at $r = -.400$, $r_s = .450$ after December 2020 ($r = -.401$, $r_s = -.452$), with a dip to $r = -.450$, $r_s = -.470$ between September 2021 ($r = -.386$, $r_s = -.424$) and December 2021 ($r = -.453$, $r_s = -.472$). This model is slightly varied for correlations between Total Deaths and average yearly temperature, as while there is a

negative strength spike during the early pandemic (April 2020, $r = -.249$, $r_s = -.420$), the correlation does not become insignificant—resembling a sinusoidal function with negative slope. Correlation between Total Deaths and temperature for December 2021 found r values of $r = -.358$, $r_s = -.420$.

Time-series correlation graphs for latitude show stronger (positive) r values across the entire 24-month period when values were statistically significant (Figure 2B). As with

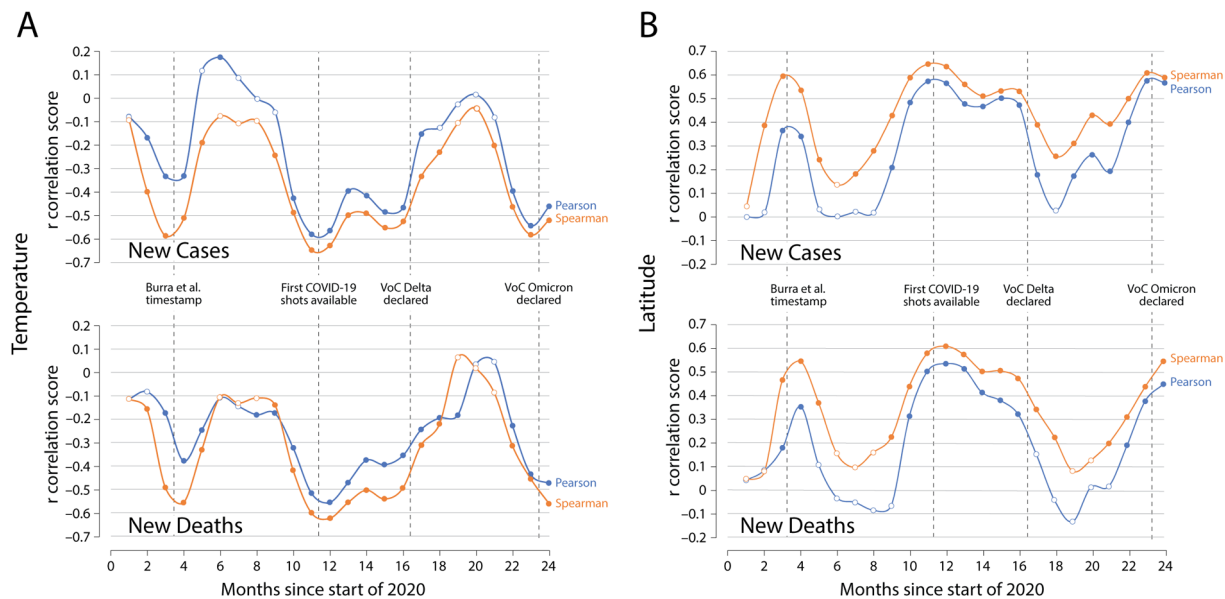


Figure 3. Time series correlation plot between New Cases (top), and New Deaths (bottom) and either average monthly temperature (A) or latitude (B). Each point represents the r coefficient of a specific month's cases or deaths and its corresponding monthly temperature. Closed symbols describe significant correlations with p -values less than .05. Average monthly temperatures of countries are difficult to compute, so a city with retrievable temperature data was selected for each country to represent this metric. Important timestamps are included in this graphic.

correlations between Total Cases or Total Deaths with temperature, there was a spike in the strength of the association during the early months of the pandemic that evens out by December 2021, which reaches correlation values of $r = .609$, $r_s = .607$ (Total Cases) and $r = .407$, $r_s = .488$ (Total Deaths) by the end of the second year.

(ii) *Correlations of New Cases or New Deaths with Temperature and Latitude.* Figure 3 shows the time series total correlation plots between New Cases and New Deaths at the end of each month and average monthly temperature and latitude. In all cases, there was an apparent cyclical trend that begins during the early months of the pandemic. Along the 24-month timeline, there were 3 noticeable negative correlation strength spikes between New Cases or New Deaths and average temperatures—a first near the beginning of the pandemic (February–April 2020), at the time we performed our first correlation analysis (Burra et al.²² timestamp), a second between October 2020 and June 2021, with the rise of Variants of Concern (VOCs) just before the global introduction of vaccines, and a final between October and December 2021, which coincided with the rise of VOC Omicron and the last month included in our time series charts. The second and third spikes reached values of $r = -.571$, $r_s = -.638$ (New Cases) and $r = -.534$, $r_s = -.634$ (New Deaths). Latitude values exhibited similar spikes. In months not included in these strength spikes, values went below $r = |.1|$ and became statistically insignificant, indicating that temperature in those months was not correlated with case incidence and deaths.

We considered the effect of summing sine waves of seasonality for the Northern and Southern Hemispheres, which are offset by 180° (as modeled in ref.¹⁵), would result in a “destructive interference” phenomenon that could cancel correlations of the more instantaneous measures of New Cases with temperature and latitude factors. However, splitting data into the 2 hemisphere-related groups, that is, 140 countries located in the Northern Hemisphere and 32 countries located on the Southern Hemisphere, revealed that global patterns were driven by the more populated and COVID-19 impacted Northern Hemisphere, which showed almost identical time-series correlation graphs (Supplemental Figure S1). The sinusoidal patterns of Figure 3 are thus driven by a more varied month-by-month response of New Cases that weakens correlations rather than destructive interference phenomena.

(iii) *Correlations of Total Cases or Total Deaths with EPI Score.* Across the study period, the correlation between Total Cases or Total Deaths variables and Yale's EPI evened out to Pearson's r values of $r = -.522$ (Total Cases) and $r = -.434$ (Total Deaths). Figure 4 depicts the corresponding time-series graphs. A significant spike in strength of correlation was observed during the early months of the pandemic (March 2020; Total Cases: $r = -.551$, $r_s = -.775$, Total Deaths: $r = -.324$, $r_s = -.094$), but steadily evened out by the end of the second year. Overall, EPI's measurement of the effort to meet UN Sustainable Development Goals was found to be correlated with case prevalence and deaths per capita of COVID-19. This is consistent with the notion that

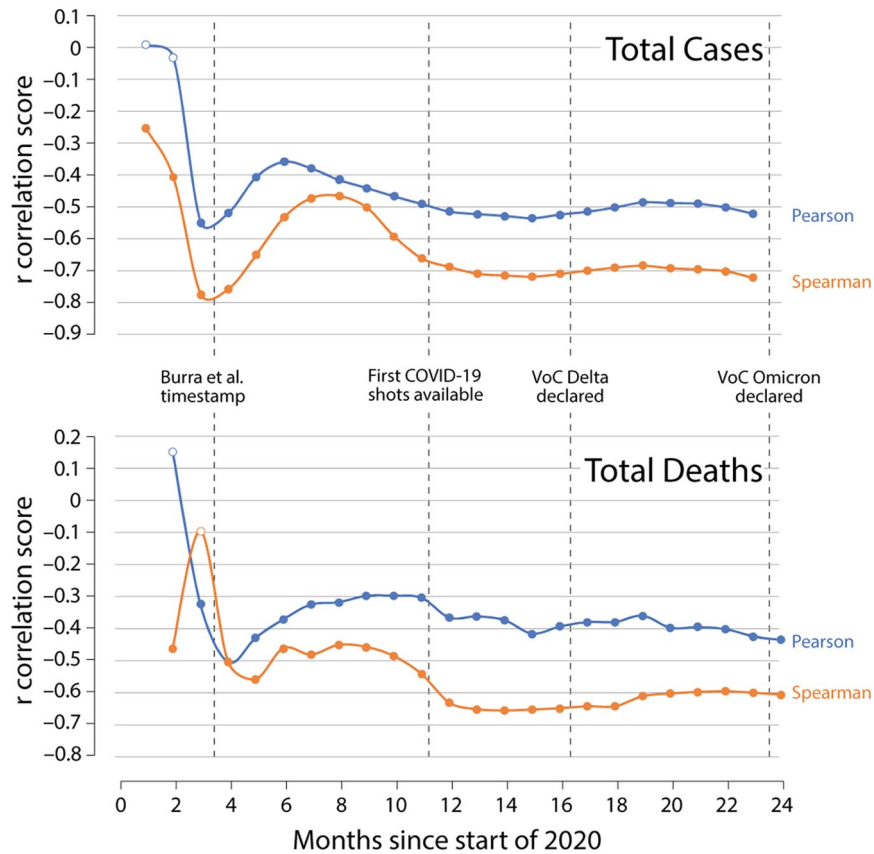


Figure 4. Time series correlation plot between Total Cases (top), Deaths (bottom) and Yale's 2020 Environmental Performance Index (EPI) scores. Each point represents the r coefficient between the Total Cases or Deaths of a specific country through each month and its EPI score. Closed symbols describe significant correlations with p -values less than .05.

sustainable development, which encompasses environmental, social, and economic dimensions, can affect overall health and well-being.

A March 2021 snapshot of categories of environmental health indicators

We found Total Cases and Total Deaths values for March 2021 were correlated with each category and performance indicator categorized by Yale's EPI score. To ensure levels of heteroscedasticity were not prevalent across categories and indicators, Total Cases and Total Deaths values were transformed by $1/\sqrt{x}$. Following this transformation, epidemiological values for March 2021 correlated with the total EPI Score at $r = -.536$, $r_s = -.718$ (Total Cases), and $r = -.464$, $r_s = -.652$ (Total Deaths). However, not all performance indicators were equally represented in the overall EPI score. Figure 5A and B depicts overall pie charts of all indicators for Total Cases and Total Deaths correlations, respectively. We highlight the most important categories and associated indicators within the "Ecosystem Vitality" and "Environmental Health" global policy objectives, which carry the largest weight as percentage of the total score and represent the largest pie slices of Figure 5:

(i) *EPI Climate Change and Pollution Metrics.* The "Climate Change" category under the EPI score had r values of $r = -.536$, $r_s = -.610$ (Total Cases) and $r = -.496$, $r_s = -.610$ (Total Deaths), while the "Pollution Emissions" category exhibited r values of $r = -.433$, $r_s = -.569$ (Total Cases) and $r = .375$, $r_s = -.507$ (Total Deaths), respectively. These categories are broken up into multiple indicators including CO_2 Growth Rates, CH_4 Growth Rates, Black Carbon Growth Rates, SO_2 Growth Rates, and NO_x Growth Rates. For each of these metrics the EPI grades according to how well a country has curbed these growth rates. For instance, a country that has significantly curbed its NO_x Growth rate receives a higher score than one who has not. We found that pollution emissions metrics correlated with COVID-19 incidence or mortality. Growth rates r values ranged from $-.509$ (CO_2) to $-.255$ (N_2O) for Total Cases and $-.481$ (CO_2) and $-.219$ (N_2O) for Total Deaths. Growth rates r_s values showed similar ranges. Surprisingly, while the GHG (Greenhouse Gas) Intensity Trend was not correlated with deaths or cases, GHG Emissions per Capita showed strong positive correlations for Total Cases ($r = .530$, $r_s = .583$) and Total Deaths ($r = .366$, $r_s = .452$). Indicators listed under Pollution Emissions showed significant negative

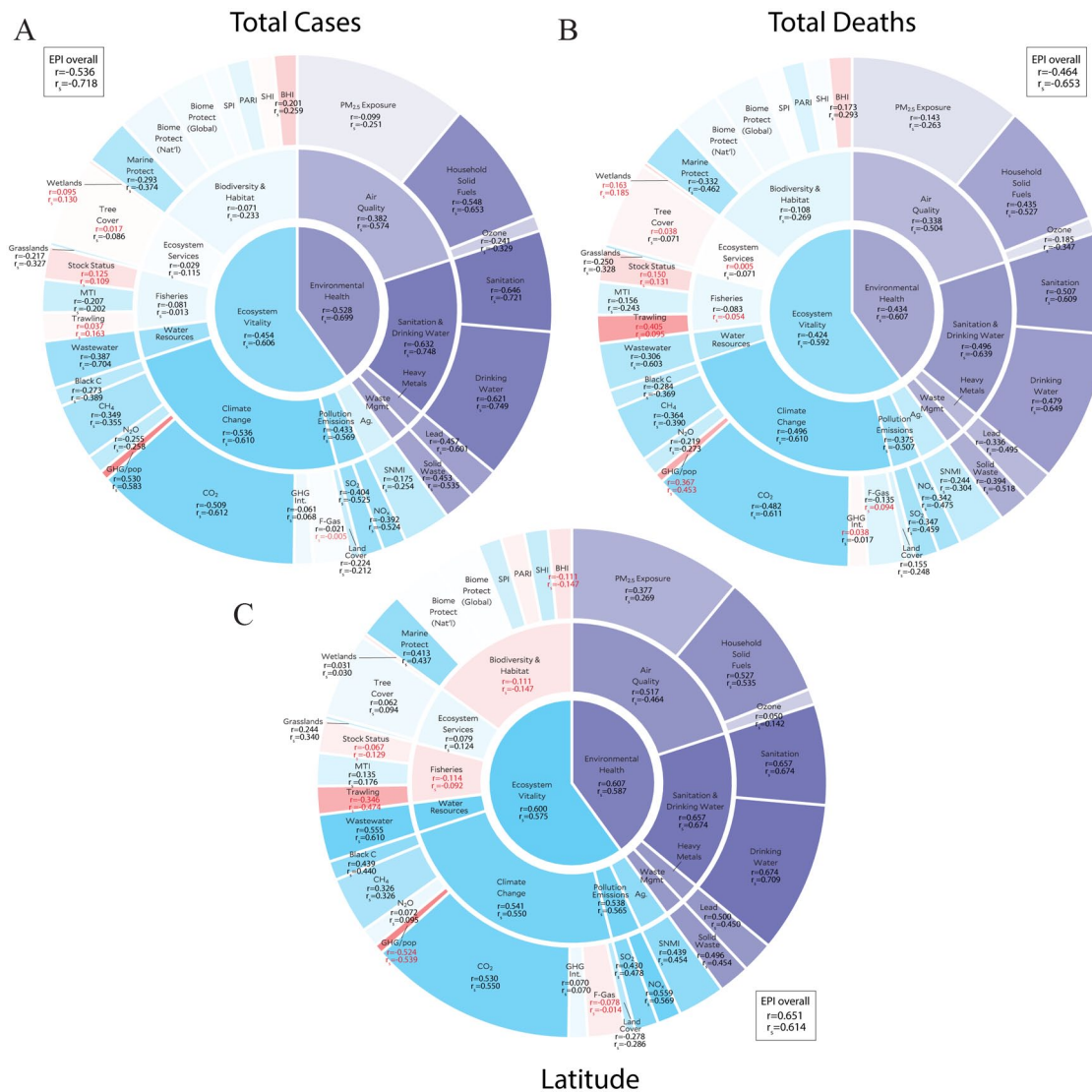


Figure 5. Tracing correlations onto a multiple pie chart describing the EPI framework. The framework organizes 32 performance indicators into 11 issue categories belonging to 2 main policy objectives, Ecosystem Vitality and Environmental Health, which are placed at the Innermost layer of the diagram. Correlations between Total Cases, Total Deaths, or latitude with EPI categories and indicators for March 2021 were traced onto the multiple pie charts. Sets with darker colors indicate stronger correlations: (A) total cases, (B) Total deaths, and (C) latitude.

correlations for SO_2 and NO_x Growth Rates for Total Cases and Total Deaths ranging $r = -.404$ to $.392$, $r_s = -.525$ to $.523$ (Total Cases) and $r = -.347$ to $.342$, $r_s = -.459$ to $.475$ (Total Deaths).

(ii) *EPI and SoGA Air Quality Metrics.* While the EPI ranks countries on country proactiveness to curve pollution, the SoGA reports population-weighted annual averages of Ozone and $\text{PM}_{2.5}$ and percentage of population using Household Solid Fuels. The EPI breaks down its “Air Quality” category into 3 indicators: $\text{PM}_{2.5}$ Exposure, Household Solid Fuels [Exposure], and Ozone (O_3) Exposure. The category showed overall correlation values of $r = -.382$, $r_s = -.574$ (Total Cases) and $r = -.338$, $r_s = -.504$ (Total Deaths). When EPI results were compared with The State of Global Air (2019) metrics for the same variables

we noted some variation. The EPI’s $\text{PM}_{2.5}$ indicator showed values of $r = -.099$, $r_s = -.251$ (Total Cases) and $r = -.142$, $r_s = -.263$ (Total Deaths), indicating insignificant to weak correlations. In contrast, SoGA’s $\text{PM}_{2.5}$ value correlations reached $r = .440$, $r_s = .509$ (Total Cases) and $r = .374$, $r_s = .499$ (Total Deaths), indicating that higher average $\text{PM}_{2.5}$ levels are more strongly correlated with an increase in COVID-19 cases and deaths than country idleness. Furthermore, The EPI’s Household Solid Fuels exposure category (representing HAP) had correlation values of $r = -.548$, $r_s = -.653$ (Total Cases) and $r = -.435$, $r_s = -.527$ (Total Deaths). The SoGA’s indicator of Household Use of Solid Fuels (representing HAP) reached higher correlation values of $r = .724$, $r_s = .608$ (Total Cases) and $r = .587$, $r_s = .485$ (Total Deaths). While the confounding factor of poverty may skew data,

EPI values showed that the use of household solid fuels were associated with case incidence and deaths. Lastly, the EPI's Ozone Exposure indicator showed negative correlation values of $r = -.241$, $r_s = -.329$ (Total Cases) and $r = -.185$, $r_s = -.346$ (Total Deaths). Conversely, SoGA's Ozone exposure by population-weighted ppb had correlation values of $r = .041$, $r_s = .035$ (Total Cases) and $r = .047$, $r_s = .124$ (Total Deaths). These results are generally considered not statistically significant, implying that Ozone Exposure was not associated with an increased or decreased risk of COVID-19 case incidence or deaths.

(iii) *Sanitation and Drinking Water Metrics*. As expected the “Sanitation and Drinking Water” category showed the most significant negative correlations, $r = -.632$, $r_s = -.748$ (Total Cases) and $r = -.496$, $r_s = -.639$ (Total Deaths). Clearly, providing necessary infrastructure for clean drinking water and sanitation affects response to any public health crisis, including that posed by the COVID-19 pandemic.

Correlation of EPI categories and performance indicators with latitude

We found the EPI score and its categories and performance indicators correlated with latitude (Figure 5C). In general, the strengths of correlations were higher than those with Total Cases and Total Deaths, as visualized by color shades proportional to Pearson's r in the multiple pie charts of the figure. For example, the correlation of the overall EPI with latitude ($r = .651$) was stronger than correlations with Total Cases ($r = -.536$) and Total Deaths ($r = -.464$). Similar patterns can be observed with the main categories of Ecosystem Vitality and Environmental Health, where the correlation with latitude ($r = .600$ and $r = .607$, respectively) was stronger than correlations with Total cases ($r = -.454$ and $r = -.528$) and Total Deaths ($r = -.424$ and $r = -.434$). Similar patterns were found with most performance indicators, with notable exceptions (eg, Biodiversity & Habitat, Fisheries). This indicates an entangled relationship exist between factors in global change analyses. We note that Spearman's r_s , which measures monotonic relationships that may not be linear, did not show differences in correlation strengths of EPI with latitude and epidemiological variables. This suggests the linear regression model may be inadequate or exhibits “omitted-variable bias,” prompting a more in-depth analysis of EPI performance indicators and the disease we will report elsewhere.

Discussion

Previous studies have hinted to the seasonal nature of COVID-19 cases and mortality (reviewed in ref. 13). However, most of these studies were performed in one or a few countries, and over short periods of time. Our study now consolidates many geographical (temperature, latitude) and epidemiological (cases and deaths) data points obtained from 171 countries situated

at different latitudes over a period of 24 months. Results show moderate to strong correlations between incidence of COVID-19 cases and deaths and air temperature and latitude. Countries with colder climates (higher latitudes) had an increased case incidence. Our study also reveals cyclical patterns in temperature and latitude variables across both COVID-19 case incidence and mortality data. An apparent bimodal trend is observed in incidence over temperature correlations—strengthening in correlation during colder months. Note that early pandemic fluctuations of COVID-19 can be explained through viral models. Caetano-Anollés et al.¹³ used SARS-CoV-2 variant genetic makeup to propose 3 phases of the disease: Phase 1, in which the disease focuses on viral survival by balancing protein flexibility and rigidity; Phase 2, which involves responding to environmental changes such as seasonal variations in temperature (for instance, incoming winter in the Northern Hemisphere); and Phase 3, which involves immune escape due to widespread vaccination. These phases of viral population structuring may help explain early correlative scores, and gradual strengthening after the first 12 months of COVID-19 prevalence.

COVID-19 seasonal effects we observe are evident regardless of the immunity of the population, behavioral changes, and the periodic appearance of new variants with higher rates of infectivity and transmissibility.³ This is particularly noteworthy because the dynamic of the pandemic has been differentially impacted by the wide range of elimination and mitigation policies and efforts applied throughout the world.^{41,42} Elimination approaches of the “zero-COVID-19” type involved rapid containment and maximum prevention through for example cancelation of gatherings, quarantines, and closing of borders between countries or states (eg, Australia). The goal was to reduce the impact of the disease to negligible levels. In contrast, mitigation approaches used relaxed measures to curb viral transmission through for example the use of masks, social distancing, and vaccination. Similarly, late vaccine introduction, vaccines with different effectiveness and responses to variants, and different vaccination rates had differential impacts on epidemiological variables in different countries.^{3,43} While continuous monitoring of these factors is critical to understand how SARS-CoV-2 may evolve and affect the world population as the disease transitions into endemicity, global seasonal behavior patterns uncovered by correlations in our study thus appears unresponsive to local epidemiological effects of immunity, behavior and viral diversity.

Sustained correlation in Total Cases and Total Deaths data along with noticeable patterns in New Cases and New Deaths data suggest countries with colder average climates (higher latitudes) fared worse as COVID-19 situated, and outlines how COVID-19 cases and deaths could become more predictable in future seasonal change. Several factors could explain why lower temperatures increase COVID-19 incidence, including those modulated by the environment, physiology,

and behavior delimiting the ‘triangle of viral persistence’.¹³ It is therefore important to understand drivers of correlation between temperature and higher incidence. The seasonal transmissibility and infectivity of SARS-CoV-2 has been shown to be correlated to the prevalence of mutations in molecular sensor regions of the spike protein, which could help the virus evade physiological responses of the host.¹³ These results suggest a physiological driver exists in the molecular structure of a crucial viral protein. However, other physiological drivers may be also factors of importance, including those involved in population immunity. In that respect, once a larger portion of the population has immunity to the prevailing variants, we expect seasonal behavior to become more predictable. This could result in better mitigation efforts, for example in focusing vaccination efforts before winter, such as in the case of influenza. While this paper provides evidence in support of seasonality through time series correlations, more research that clearly shows the driver of lower temperature and higher incidences of COVID-19 is needed.

Our study also uncovers correlations between environmental health indicators and case incidence and mortality of COVID-19. Indicators included air quality and pollution emission metrics. These correlations are particularly important, as these are environmental variables that can be modified by policy. In sharp contrast, temperature or latitude cannot be acted upon. The Yale’s 2020 EPI is a summary of the state of sustainability around the world.³⁸ The EPI ranks countries on a scale of 0 to 100 by how close they are to meeting the targets of the UN Sustainable Development Goals and how well they are addressing global environmental problems. The EPI further breaks down its overall ranking into 32 performance indicators cataloged into 11 categories. Countries that are on-track to meet these goals are awarded a higher score, and those that are not on-track or have regressed receive a lower score. As such, the EPI Index is not a measure of actual prevalence of certain indicators in each country, but instead an indicator of country proactivity to improve public and environmental health. While not all categories are related to viral or bacterial spread, the Air Quality (PM_{2.5} Exposure, Household Solid Fuels, O₃ Exposure), Pollution Emissions (SO₂, NO_x Growth Rates), and Climate Change (CO₂, N₂O, CH₄, Black Carbon Growth Rates) categories are known to be correlated with disease prevalence and outcome^{26,27} and therefore represent potential targets of correlation with COVID-19 incidence and mortality. EPI scores are based off of manipulation of datasets that satisfy the inclusion criteria of the Yale’s EPI report. This criterion includes evaluation of a dataset relevance, methodology, verification (peer reviewed or third-party auditing and confirmation), temporal completeness, and recency. Inclusion of a variety of datasets allows the EPI to evaluate as much as it can about an included country. The State of Global Air (SoGA) Report³⁴ is produced yearly by the Health Effects Institute and Institute for Health Metrics and Evaluations. The SoGA

reports average and median national levels of Ozone Exposure, PM_{2.5} Exposure, and national percentages of people using Household Solid Fuels. In contrast to the EPI, SoGA does not transform or index its values based on other targets. Instead, SoGA reports raw data in units it is commonly measured in. Using both the EPI and SoGA to quantify not only the prevalence of air pollutants in each country, but also the proactivity of countries to improve their air quality allows for the bilateral analysis of Air Quality effects on COVID-19.

Remarkably, we show a high correlation exists between all environmental variables studied, particularly HAP, Climate Change, and Pollution Emission metrics. These metrics could be used as a foundation of environmental policy changes needed at a country level for improving public health. Poor air quality and higher incidence of deaths from COVID-19 was previously demonstrated in the US.⁴⁴ This effect was independent of other risk factors such as socioeconomic status, access to health care, and pre-existing medical conditions. Our study shows that environmental stressors contribute to COVID-19 severity observed on a global scale. HAP for example was strongly associated with severity of COVID-19, exhibiting highly significant correlations of both EPI and SoGA scores. In contrast, while EPI correlations between cases/deaths and PM_{2.5} were weak, SoGA correlations were 2 to 4 times stronger. PM_{2.5} score variations suggest differences between country proactivity scores and real PM_{2.5} data.

While CO₂, SO₂ and other compounds are present in PM_{2.5} and contribute to HAP pollutants, they are measured separately as part of the Climate Change and Pollution Emission categories under the EPI scheme. The Climate Change category correlation was substantial, at $r = -.536$, $r_s = -.610$ (Total Cases), and $r = -.496$, $r_s = -.610$ (Total Deaths). Pollution Emission correlation was also significant, at $r = -.433$, $r_s = -.569$ (cases), and $r = -.375$, $r_s = -.506$ (deaths). Multiple indicators (CO₂, CH₄, SO₂, NO_x) within these categories contribute to respiratory co-morbidities both in confined spaces and as ambient pollutants. When modeling disease prevention nationally and globally, these indicators can help evaluate the long-term effects of pollutants.

One remarkable but complicating finding was our observation that EPI performance indicators of global change were correlated with latitude. While to our knowledge EPI-latitude associations have not been previously reported, EPI indicators are known to exhibit a strong positive correlation with gross domestic product (GDP) per capita, which measures the monetary value of final goods and services of a country and serves as an indicator of wealth.³⁸ It is also well known that economic development and latitude are positively correlated and that such correlation evolved from a negative correlation that existed 500 years ago.⁴⁵⁻⁴⁷ An economic and physiological developmental model explained this latitude gradient by claiming metabolic costs of fertility during human adaptation to cold temperatures allowed a “reversal of fortune” process of

sustained growth that eventually led to increasing long-term economic development at higher latitudes.⁴⁷ This model is aligned with links between available energy needed to produce goods and services (measured as “energy return on investment,” EROI) and social well-being.⁴⁸ In those studies, energy indicators were found to be highly correlated with indicators of quality of life such as the Human Development Index (HDI), health expenditures, and access to water. Remarkably, latitudinal gradients of indicators of quality of life have been shown across the globe.⁴⁹ Creativity, aggression, and happiness correlated with latitude as well as climatic remote (thermal demands, steady rain) and proximate (pathogen prevalence, national wealth) predictors. Recently, other measures of quality of life have been shown to be also correlated with latitude, including the HDI and the world happiness score (WHS).⁵⁰ Thus cultural and psychological diversity in populations unfold latitudinally on Earth impacting wealth and happiness. We contend that these factors also impact global change as well as COVID-19 predictors, which may well be signaling another “reversal of fortune” in latitudinal patterns.

We end by noting that correlation does not imply causation. Neither Pearson’s product-moment correlation nor Spearman’s rank-order correlation indicate a relationship between 2 variables is causal (necessary, sufficient, and contributing to a cause), only that a relationship possibly exists. As such, we use these correlations alongside graphs and additional data to provide a clearer picture of our exploration. They should be considered first steps in the search for mechanistic links between both seasonal and global change effects on the disease that could be operating at environmental, physiological, and behavioral levels. In fact, 5 criteria should be fulfilled to establish causality.⁵¹ First, there must be an empirical association between independent and dependent variables established with correlation analyses, which we have shown exists in our study. Without such association there cannot be a causal relationship. Second, variations in the independent variable must precede variations in the dependent variable. This “temporal priority” criterion guarantees that the cause precedes its effects. Third, association must not be spurious. The association cannot be caused by some other confounding variable. One example of possible confounding variables for Total Cases is the number of COVID-19 tests used to report cases of the disease (testing rate), which in a previous study failed to correlate with temperature and latitude.²² In turn, confounding variables for Total Deaths are demographic factors (eg, age, sex, income), which have not been explored in this study. In terms of global change parameters, we did find an important association between EPI performance indicators and latitude, suggesting a need to disentangle the seasonal and global change effects on the disease. Fourth, there should be a mechanism that connects independent and dependent variables. For example, latitude is responsible for seasonal temperature variations because of Earth’s tilted axis relative to the plane of its orbit, which could in turn affect the infectivity and survival of a virus seeking to complete its life

cycle. This causal chain has been used to explain periodicities of influenza⁵² and could similarly apply to COVID-19. Finally, there is a contextual framework surrounding causation that strengthens the cause-effect relationship. For example, elevation decreases temperature but is not correlated with Total Cases or Total Deaths of COVID-19 worldwide.¹³ Similarly, merging latitude and temperature effects from Southern and Northern Hemispheres fail to offset correlations but strengthens our study. To conclude, there is still many steps to complete and much to learn about the causal factors underlying epidemic calendars and global change dynamics.

Conclusions

Our time-series correlation study highlights the central role that seasonal and global change effects plays in the ongoing COVID-19 pandemic. Whereas the dynamic of the pandemic has changed often with the rise of new variants and diverse elimination and mitigation efforts, the high association of factors studied in this work with COVID-19 incidence and mortality has remained constant. Foundations of responsible environmental policies at a country level must therefore be established to reduce the effect of biotic and environmental stressors and improve public health. We stress: *countries that go against the health of the planet affect public health in general.*⁵³

Acknowledgements

COVID-19 research in the laboratory of G.C.-A is supported by the Office of Research and Office of International Programs in the College of Agricultural, Consumer and Environmental Sciences at the University of Illinois at Urbana-Champaign and New Frontiers Initiative, National Center for Supercomputing Applications (NCSA).

Author Contributions

N.H. collected epidemiological and geographical data and analyzed results. N.H. and G.C.A. designed the project, contributed to data analysis and interpretation, and wrote, edited and approved the manuscript.

Supplemental Material

Supplemental material for this article is available online.

REFERENCES

1. Wu F, Zhao S, Yu B, et al. A new coronavirus associated with human respiratory disease in China. *Nature*. 2020;579:265-269.
2. John Hopkins Coronavirus Resource Center. *COVID-19 Dashboard*, John Hopkins University & Medicine; 2023. Accessed February 15, 2023. <https://coronavirus.jhu.edu/map.html>
3. Telenti A, Hodcroft EB, Robertson DL. The evolution and biology of SARS-CoV-2 variants. *Cold Spring Harb Perspect Med*. 2022;12:a041390.
4. Wiersinga WJ, Rhodes A, Cheng AC, Peacock SJ, Prescott HC. Pathophysiology, transmission, diagnosis, and treatment of coronavirus disease 2019 (COVID-19). *JAMA*. 2020;324:782-793.
5. Abdelrahman Z, Li M, Wang X. Comparative review of SARS-COV-2, SARS-COV, MERS-COV, and influenza A respiratory viruses. *Front Immunol*. 2020; 11:552909. doi:10.3389/fimmu.2020.552909
6. Riddell S, Goldie S, Hill A, Eagles D, Drew TW. The effect of temperature on persistence of SARS-COV-2 on common surfaces. *BMC Virology J*. 2020;17(1): 145. doi:10.1186/s12985-020-01418-7

7. Molina-Mora JA, González A, Jiménez-Morgan S, et al. Clinical profiles at the time of diagnosis of SARS-CoV-2 infection in Costa Rica during the pre-vaccination period using a machine learning approach. *Phenomics*. 2022;2:312-322. doi:10.1007/s43657-022-00058-x
8. Oran DP, Topol EJ. The proportion of SARS-CoV-2 infections that are asymptomatic: a systematic review. *Ann Intern Med*. 2021;174(5):M20655-6976662. doi:10.7326/M20-6976
9. Martinez ME. The calendar of epidemics: seasonal cycles of infectious diseases. *PLoS Pathog*. 2018;14:e1007327.
10. Brownlee J. VI. An investigation into the periodicity of measles epidemics in London from 1703 to the present day by the method of the periodogram. *Philos Trans R Soc Lond B Containing Pap Biol Character (1896-1934)*. 1918;208:225-250. doi:10.1098/rstb.1918.0006
11. Soper HE. The interpretation of periodicity in disease prevalence. *J R Stat Soc*. 1929;92:34-73.
12. Grenfell B, Björnstad O. Sexually transmitted diseases: epidemic cycling and immunity. *Nature*. 2005;433(7024):366-367. https://doi.org/10.1038_433366a
13. Caetano-Anollés G, Hernandez N, Mughal F, Tomaszewski T, Caetano-Anollés G. The seasonal behavior of COVID-19 and it galectin-like culprit of the viral spike. *Methods Microbiol*. 2021;50:27-81.
14. Nickbakhsh S, Ho A, Marques DFP, McMenamin J, Gunson RN, Murcia PR. Epidemiology of seasonal coronaviruses: establishing the context for the emergence of coronavirus disease 2019. *J Infect Dis*. 2020;222:17-25.
15. Matson MJ, Yinda CK, Seifert SN, et al. Effect of environmental conditions on SARS-COV-2 stability in human nasal mucus and sputum. *Emerg Infect Dis*. 2020;26:2276-2278.
16. Chan KH, Peiris JS, Lam SY, Poon LL, Yuen KY, Seto WH. The effects of temperature and relative humidity on the viability of the SARS coronavirus. *Adv Virol*. 2011;2011:1734690-1734697. doi:10.1155/2011/734690.
17. Prata DN, Rodrigues W, Bermejo PH. Temperature significantly changes COVID-19 transmission in (sub)tropical cities of Brazil. *Sci Total Environ*. 2020;729:138862. doi:10.1016/j.scitotenv.2020.138862
18. Wu Y, Jing W, Liu J, et al. Effects of temperature and humidity on the daily new cases and new deaths of COVID-19 in 166 countries. *Sci Total Environ*. 2020;729:139051.
19. Menebo MM. Temperature and precipitation associate with COVID-19 new daily cases: A correlation study between weather and covid-19 pandemic in Oslo, Norway. *Sci Total Environ*. 2020;737:139659.
20. Xie J, Zhu Y. Association between ambient temperature and COVID-19 infection in 122 cities from China. *Sci Total Environ*. 2020;724:138201.
21. Kassem AZE. Does temperature affect COVID-19 transmission? *Front Public Health*. 2020;8:554964.
22. Burra P, Soto-Díaz K, Chalen I, Gonzalez-Ricon RJ, Istanto D, Caetano-Anollés G. Temperature and latitude correlate with SARS-COV-2 epidemiological variables but not with genomic change worldwide. *Evol Bioinform*. 2021;17:1176934321989695.
23. Liu X, Huang J, Li C, et al. The role of seasonality in the spread of COVID-19 pandemic. *Environ Res*. 2021;195:110874.
24. López-Feldman A, Heres D, Marquez-Padilla F. Air pollution exposure and COVID-19: A look at mortality in Mexico City using individual-level data. *Sci Total Environ*. 2021;756:143929.
25. Kelly FJ, Fussell JC. Air pollution and airway disease. *Clin Exp Allergy*. 2011;41:1059-1071.
26. Streimikiene D. Environmental indicators for the assessment of quality of life. *Intellect Econ*. 2015;9:67-79.
27. Xing YF, Xu YH, Shi MH, Lian YX. The impact of PM2.5 on the human respiratory system. *J Thorac Dis*. 2016;8:E69-E74.
28. Karan A, Ali K, Teelucksingh S, Sakhamuri S. The impact of air pollution on the incidence and mortality of COVID-19. *Glob Health Res Policy*. 2020;5:39.
29. World Health Organization. *Household Air Pollution and Health*. World Health Organization; 2021, September 22. Accessed March 6, 2022. https://www.who.int/news-room/fact-sheets/detail/household-air-pollution-and-health
30. Shupler M, Godwin W, Frostad J, Gustafson P, Arku RE, Brauer M. Global estimation of exposure to fine particulate matter (PM2.5) from Household Air Pollution. *Environ Int*. 2018;120:354-363.
31. Dominici F, Wang Y, Correia AW, Ezzati M, Pope Ca 3rd, Dockery D. W. Chemical composition of fine particulate matter and life expectancy. *Epidemiology*. 2015;26(4):556-564. doi:10.1097/ede.0000000000000297
32. Hystad P, Duong M, Brauer M, et al. Health effects of household solid fuel use: findings from 11 countries within the prospective urban and rural epidemiology study. *Environ Health Perspect*. 2019;127:057003.
33. Chan KH, Kurmi OP, Bennett DA, et al. Solid fuel use and risks of respiratory diseases. A cohort study of 280,000 Chinese never-smokers. *Am J Respir Crit Care Med*. 2019;199:352-361.
34. Health Effects Institute. 2019. State of Global Air 2019. www.stateofglobalair.org
35. D'Amato G, Cecchi L, D'Amato M, Annesi-Maesano I. Climate change and respiratory diseases. *Eur Respir Rev*. 2014;23:161-169.
36. Wang M, Aaron CP, Madrigano J, et al. Association between long-term exposure to ambient air pollution and change in quantitatively assessed emphysema and lung function. *JAMA*. 2019;322:546-556.
37. Wilson A, Reich BJ, Nolte CG, Spero TL, Hubbell B, Rappold AG. Climate change impacts on projections of excess mortality at 2030 using spatially-varying ozone-temperature risk surfaces. *J Expo Sci Environ Epidemiol*. 2016;27:118-124.
38. Wendling ZA, Emerson JW, de Sherbinin A, Esty DC, et al. *2020 Environmental Performance Index*. Yale Center for Environmental Law & Policy; 2020. https://epi.yale.edu/
39. Schober P, Boer C, Schwarte LA. Correlation coefficients. *Anesth Analg*. 2018;126:1763-1768.
40. Global Change Data Lab. *Our World in Data*; 2023. https://ourworldindata.org
41. Oliu-Barton M, Pradelski BSR, Algan Y, et al. Elimination versus mitigation of SARS-CoV-2 in the presence of effective vaccines. *Lancet Glob Health*. 2022;10:e142-e147. doi:10.1016/S2214-109X(21)00494-0
42. Su Z, Cheshmehzangi A, McDonnell D, et al. The advantages of the Zero-COVID-19 strategy. *Int J Environ Res Public Health*. 2022;19:8767.
43. McLean G, Kamil J, Lee B, et al. The impact of evolving SARS-CoV-2 mutations and variants on COVID-19 vaccines. *mBio*. 2022;13:1-24.
44. Wu X, Nethery RC, Sabath MB, Braun D, Dominici F. Air pollution and COVID-19 mortality in the United States: strengths and limitations of an ecological regression analysis. *Sci Adv*. 2020;6:eabd4049.
45. Acemoglu D, Johnson S, Robinson JA. Reversal of fortune: geography and institutions in the making of the modern world income distribution. *Q J Econ*. 2002;117:1231-1294.
46. Ashraf Q, Galor O. Dynamics and stagnation in the Malthusian epoch. *Am Econ Rev*. 2011;101:2003-2041.
47. Dalgaard C-J, Strulik H. Physiological constraints and comparative economic development. CESifo Working Paper, No. 6794, 2017. Center for Economic Studies & Ifo Institute (CESifo), Munich.
48. Lambert JG, Hall CAS, Balogh S, Gupta A, Arnold M. Energy, EROI and quality of life. *Energy Policy*. 2014;64:153-167.
49. Van de Vliert E, Van Lange PAM. Latitudinal psychology: an ecological perspective on creativity, aggression, happiness, and beyond. *Perspect Psychol Sci*. 2019;14:860-884.
50. Pawliczek A, Kurowska-Pysz J, Smilnak R. Relation between globe latitude and the quality of life: Insights for public policy management. *Sustainability*. 2022;14:1461.
51. Chambliss DF, Schutt RK. *Making Sense of the Social World*. 4th ed. SAGE Publications; 2012.
52. Deyle ER, Maher MC, Hernandez RD, Basu S, Sugihara G. Global environmental drivers of influenza. *Proc Natl Acad Sci USA*. 2016;113:13081-13086.
53. The Lancet Public Health. No public health without planetary health. *Lancet*. 2022;7(4):e291. doi:10.1016/S2468-2667(22)00068-8