

## Research Articles

# A statistical assessment of association between meteorological parameters and COVID-19 pandemic in 10 countries

Shola Adeyemi<sup>1</sup> , Usame Yakutcan<sup>2</sup> , Eren Demir<sup>2</sup> <sup>1</sup> Bohemian Smartlytics Ltd, Haverhill, UK, <sup>2</sup> University of Hertfordshire, Hertfordshire Business School, Hatfield, UK

Keywords: meteorological parameters, stringency index, granger causality, covid-19

<https://doi.org/10.29392/001c.13693>

---

## Journal of Global Health Reports

Vol. 4, 2020

---

### Background

Eleven out of 13 published articles reported temperature and humidity as factors that could reduce the daily confirmed COVID-19 cases among many other findings. However, there are significant caveats, related to statistical assumptions and the spatial-temporal nature of the data.

### Methods

Associative and causative analyses of data was conducted for 10 countries representing 6 continents of the world, with data obtained between January 22, 2020 to April 30, 2020. Daily confirmed cases, number of deaths, recovered cases, lockdown stringency index, and several meteorological factors are considered. Also, a Granger-Causality test was performed to check if any COVID-19 outcomes are influenced by itself and not by any or combination of maximum temperature, humidity, wind speed and stringency index.

### Results

Most of the associations reported in the literature, between meteorological parameters and COVID-19 pandemic are weak evidence, need to be interpreted with caution, as most of these articles neglected the temporal spatial nature of the data. Based on our findings, most of the correlations no matter which coefficient is used are mostly and strictly between -0.5 and 0.5, and these are weak correlations. An interesting finding is the correlation between stringency and each of the COVID-19 outcomes, the strongest being between stringency and confirmed cases, 0.80 (0.78, 0.82)  $P < .0001$ . Similarly, wind speed is weakly associated with recovery rate, 0.22 (0.16, 0.28)  $P < .0001$ . Lastly, the Granger-Causality test of no dependencies was accepted at  $P = 0.1593$ , suggesting independence among the parameters.

### Conclusions

Although many articles reported association between meteorological parameters and COVID-19, they mainly lack strong evidence and clear interpretation of the statistical results (e.g. underlying assumption, confidence intervals, a clear hypothesis). Our findings showed that, without effective control measures, strong outbreaks are likely in more windy climates and summer weather, humidity or warmer temperature will not substantially limit pandemic growth.

If severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) conforms to the seasonal pattern of other respiratory infections, then the arrival of spring should help reduce the spread of the disease.<sup>1-3</sup> This assumption has prompted researchers around the world to investigate whether meteorological factors (e.g. temperature and humidity) had any effect on daily incidence of coronavirus disease (COVID-19) (e.g. death and confirmed cases). Table 1 summarises the findings from all the published articles so far, where 11 out of the 13 articles reported at least one meteorological factor having some form of an association on daily confirmed cases (or deaths).

Temperature and relative humidity were the two factors that are mostly found to be statistically significant. For instance, according to Islam et al.,<sup>4</sup> there is an inverse relationship between temperature and humidity and the incidence of COVID-19, suggesting that colder and drier environments are more favourable condition for virus survival.

Similar findings were reported by.<sup>5-10</sup> However, Yao et al.<sup>11</sup> found no association of COVID-19 transmission with temperature or UV radiation, whereas Xie and Zhu<sup>12</sup> reported positive relationship between mean temperature and the number of confirmed cases with a threshold of 3°C, but no evidence supporting that confirmed cases of the virus declines when the weather becomes warmer. Bannister-Tyrrell et al.<sup>13</sup> concludes that there is a possible seasonal variability, but states that it does not imply temperature alone is a primary driver. Clearly, there is no consensus from the scientific community on this pressing issue.

The methodologies used to establish the link can be divided into two, simple correlation analysis and regression analysis. The latter approach includes multiple linear regression, generalized linear model (GLM), and the generalized additive model (GAM), where COVID-19 outcome is used as the dependent variable and meteorological factors as explanatory (or independent) variable (see Table 1). Lag

effects of weather conditions, such as 14 days was embedded in the form of covariates.

However, there are significant caveats in all these studies, mostly related to methodological assumptions and the data, namely i) the assumption that the data is normally distributed, and using Pearson correlation, when in fact after visual inspection, the data shows huge deviation from normality, ii) limitation in time and location, for instance, examining a single country is insufficient, whereas worldwide could easily distort the analysis, iii) correlation looks for linear relationship between two variables, however, on visual inspection there is little or no linearity in the relationships, and iv) the use of Spearman's correlation has a bias towards zero.

In all studies but Ma et al.,<sup>5</sup> the confirmed daily cases were used as COVID-19 outcome, which heavily relies on countries testing capacity. Germany and South Korea has led the way, and Spain, United Kingdom (UK) and South Africa, by contrast, have limited their testing to the very sick or have been constrained by a shortage of testing kits. Therefore, daily confirmed cases may not tell the full story, thus additional outcomes, such as recovery and daily deaths should be considered, to capture the true effect of meteorological factors. In addition, stringency index (SI), a score ranging from 0 – 100, should also be part of the study, to assess the impact on the spread of the disease measured by the reproduction number R.<sup>17</sup> A lower R number is better, ideally lower than 1, meaning that each infected person could potentially pass it to less than one person on average. SI scores a governments response to the pandemic in terms of reducing mobility of its population, such as a complete or a partial lockdown, e.g. closure of businesses, schools, and transport links.<sup>18</sup>

## METHODS

### DATA SOURCES

Data, including COVID-19, meteorological, SI, welfare level, and health care capacity of countries, were obtained from 22nd January 2020 to 30th April 2020 for 10 countries to represent a worldwide approach. Australia, Brazil, Canada, Germany, Italy, New Zealand, Singapore, South Africa, Sweden, and UK were selected for the analysis to capture at least one destination in each continent (except Antarctica).

Each government and its population have a unique strategy/reaction against the outbreak, e.g. strictness of lockdown, obedience to rules, wearing of masks and closing borders. For instance, a severe lockdown in Italy and United Kingdom, whereas Sweden preferred a much more relaxed approach. Also, the data availability, accuracy along with democracy and freedom are other important issues, e.g., coding COVID-19 related deaths vary a lot between countries. Thus, examining a single country to make a generic deduction is biased and not accurate. Moreover, the inclusion of all countries (rather than being selective) could easily distort the analysis. Therefore, by considering these issues, we chose the listed 10 countries to reflect the variance in the indicators affecting the spread of the disease.

SI, confirmed COVID-19 cases, COVID-19 related deaths, recovered cases as well as maximum/minimum temperature, humidity, and wind speed, were our primary factors of interest (daily basis) as their associations are covered partially in the literature. In addition, number of hospital beds per 1000 population,<sup>19</sup> hospital occupancy rates and US dollars per capital spend on health<sup>20</sup> was used as a measure of each country's health systems preparedness for

COVID-19. Population density per square kilometre is also considered as high population density could make people more vulnerable because of the possibility of frequent social contacts.

COVID-19 figures were drawn from the rich database developed by the Johns Hopkins University (data available at <https://github.com/CSSEGISandData/COVID-19>). Meteorological data were obtained from Dark Sky (data available at <https://github.com/imantsm>) and Visual Crossing (data available at <https://www.visualcrossing.com/weather-data>). Lastly, stringency index data, which scales the government responses to the outbreak, were taken from the dataset constituted by the Blavatnik School of Government, University of Oxford (data available at <https://github.com/OxCGRT/covid-policy-tracker>).

### STATISTICAL ANALYSIS

Descriptive statistics were used to summarize the data explained as presented in the summary statistics of the data (Table 2). For Australia and Canada, because the temperature can vary a lot among the different provinces/regions, we have assumed minimum/maximum of all regional minima/maxima, average wind speed and humidity. Table 2 shows high variability in confirmed cases, recovery rates and death amongst the 10 countries. It also shows that these parameters are far from being normally distributed, as the mean and median are far apart, except for maximum temperature and hospital occupancy, which is not surprising as the WHO benchmark on optimal occupancy is 80%.<sup>20</sup>

Tasks associated with the analysis of temporal datasets typically focus on the evolution of other (dependent) variables with respect to time: identifying trends and recurring patterns, establishing correlations, and possibly predicting the future based on past and current behaviour. The aim of this study is to conduct an exploratory time series analysis to identify and interpreting the relationships between COVID-19 outcomes and weather parameters based on trends, recurring patterns, autocorrelations, cross correlations, cointegration and identifying causality. Prediction of future COVID-19 outcomes is beyond the scope of this study.

Using a frequency domain, spectral decomposition with no smoothing, we generate periodograms, autocorrelation functions and cross correlations. Then, using a vector autoregressive model, causality was tested among the weather parameters and COVID-19 outcomes using Granger Wald test.

## RESULTS

Maximum temperature is reported to have some relationship with COVID-19 outcomes, hence selected accordingly, as opposed to minimum temperature. We plotted graphs for the temporal pattern of COVID-19 figures against daily maximum temperature for each country. The graphs and their more detailed interpretation are presented in the **Online Supplementary Document** (see Figure S1 and Figure S2). The plots showed that the figures (daily number of recoveries and daily mortality) had a similar pattern with daily maximum temperature in most countries. These patterns and relationship need to be investigated in a further detail by statistical approaches along with other variables.

We explore the raw correlation between confirmed cases, recovery rates, number of death and the weather and country characteristics as stated above. Table 3 below presents the correlation estimate and associated confidence interval

**Table 1: A summary of current studies investigating the effect of meteorological factors on the incidence of COVID-19**

Author(s)	COVID-19 outcome	Meteorological factors	Methodology	Was there any meteorological effect on COVID-19 outcome?	The meteorological factors found to be significant	Location
Ma et al. <sup>5</sup>	Number of daily deaths	<ul style="list-style-type: none"> <li>• Diurnal temperature range</li> <li>• Daily average temperature</li> <li>• Relative humidity</li> <li>• Air pollutants</li> </ul>	Generalized Additive Model	Yes	<ul style="list-style-type: none"> <li>• Diurnal temperature range is positively associated with Covid-19 outcome</li> <li>• Relative humidity and ambient temperature are negatively associated</li> </ul>	Wuhan, China
Chen et al. <sup>6</sup>	Confirmed daily cases	<ul style="list-style-type: none"> <li>• Air temperature</li> <li>• Relative humidity</li> <li>• Wind speed</li> <li>• Visibility</li> </ul>	Generalized Linear Model	Some evidence when factors are combined	<ul style="list-style-type: none"> <li>• Air temperature</li> <li>• Relative humidity</li> <li>• Wind speed</li> <li>• Visibility</li> </ul>	Worldwide
Bashir et al. <sup>14</sup>	Confirmed daily cases	<ul style="list-style-type: none"> <li>• Average daily temperature</li> <li>• Min &amp; Max temperature</li> <li>• Rainfall</li> <li>• Relative humidity</li> <li>• Wind speed</li> <li>• Air quality</li> </ul>	Kendall & Spearman Rank Correlation	Yes	<ul style="list-style-type: none"> <li>• Average daily temperature</li> <li>• Minimum temperature</li> <li>• Air quality</li> </ul>	New York (USA)
Islam et al. <sup>4</sup>	Confirmed daily cases	<ul style="list-style-type: none"> <li>• Temperature</li> <li>• Relative humidity</li> <li>• Wind speed</li> <li>• Ultraviolet index</li> </ul>	Multilevel mixed effects negative binomial regression	Yes	<ul style="list-style-type: none"> <li>• Temperature</li> <li>• Relative humidity</li> <li>• Wind speed</li> </ul>	Worldwide
Yao et al. <sup>11</sup>	Confirmed daily cases	<ul style="list-style-type: none"> <li>• Avg., min and max temperature</li> <li>• Relative humidity</li> <li>• UV radiation</li> </ul>	Descriptive analysis, correlations and chi-square test	No	N/A	China
Xie and Zhu <sup>12</sup>	Confirmed daily cases	<ul style="list-style-type: none"> <li>• Mean daily temperature</li> <li>• Relative humidity</li> <li>• Air pressure</li> <li>• Wind speed</li> </ul>	Generalized Additive Model	No	Mean temperature and confirmed cases was approximately linear in the range of < 3°C, but flat above 3°C.	China
Oliverios et al. <sup>7</sup>	Confirmed daily cases	<ul style="list-style-type: none"> <li>• Temperature</li> <li>• Relative humidity</li> <li>• Rainfall</li> <li>• Wind speed</li> </ul>	Multiple Linear Regression	Yes	<ul style="list-style-type: none"> <li>• Temperature</li> <li>• Relative humidity</li> </ul>	China

Author(s)	COVID-19 outcome	Meteorological factors	Methodology	Was there any meteorological effect on COVID-19 outcome?	The meteorological factors found to be significant	Location
Notari <sup>8</sup>	Confirmed daily cases	<ul style="list-style-type: none"> <li>• Temperature</li> <li>• Relative humidity</li> </ul>	Linear Regression	Yes	<ul style="list-style-type: none"> <li>• Temperature</li> <li>• Relative humidity</li> </ul>	Worldwide
Sajadi et al. <sup>9</sup>	Confirmed daily cases and deaths	<ul style="list-style-type: none"> <li>• Temperature</li> <li>• Relative humidity</li> <li>• Specific humidity</li> <li>• Absolute humidity</li> </ul>	Descriptive and correlational analysis	Yes	<ul style="list-style-type: none"> <li>• Temperature</li> <li>• Relative humidity</li> </ul>	Worldwide
Wang et al. <sup>10</sup>	Confirmed daily cases	<ul style="list-style-type: none"> <li>• Temperature</li> <li>• Relative humidity</li> </ul>	Linear Regression	Yes	<ul style="list-style-type: none"> <li>• High temperature</li> <li>• High humidity</li> </ul>	Worldwide
Bannister-Tyrrell et al. <sup>13</sup>	Confirmed daily cases	<ul style="list-style-type: none"> <li>• Daily temperature</li> </ul>	Regression analysis	Yes/No	Possible seasonal variability but concludes that it does not imply temperature alone is a primary driver.	Worldwide
Sahin <sup>15</sup>	Confirmed daily cases	<ul style="list-style-type: none"> <li>• Daily average temperature</li> <li>• Daily average humidity</li> <li>• Daily average wind speed</li> </ul>	Correlation Analysis	Yes	<ul style="list-style-type: none"> <li>• Temperature</li> <li>• Wind speed</li> </ul>	Turkey
Tosepu et al. <sup>16</sup>	Confirmed daily cases	<ul style="list-style-type: none"> <li>• Min, max and average daily temperature</li> <li>• Humidity</li> <li>• Rainfall</li> </ul>	Spearman-rank correlation	Yes	<ul style="list-style-type: none"> <li>• Temperature</li> </ul>	Jakarta, Indonesia

and *P*-values.

Table 3 shows the devil in the details of the correlational analysis, without the confidence intervals one would believe that most of these parameters are correlated, especially with such low *p*-values. The test for which the *p*-values are generated is whether the confidence intervals exclude zeros, so as far as this is satisfied the *P*-values could be significant. However, most of these results are very weak correlation and could even be spurious based on arguments already put forward in the introduction section. The correlations that could potentially make some sense are those between stringency and each of recovered 0.60 (0.56, 0.64) *P*<.0001, confirmed 0.80 (0.78, 0.82) *P*<.0001 and death 0.69 (0.66, 0.72) *P*<.0001.

Exploring these correlations visually reveals what we have argued. Figure 1 brings into limelight the hidden truth that portrays the weak evidence of association between meteorological parameters and COVID-19 outcomes. None of these parameters can be suggested to satisfy the assumption of normality in which the Pearson correlation coefficient relies. Also, correlation identifies how strong is the

linear relationship between the two variables, however, the plots of weather parameters and COVID-19 cases above shows very weak or no linearity in the relationships. The use of spearman's rank correlation, to correct the normality assumption when violated, is always bias towards zero.<sup>21</sup> As we can see most of the COVID-19 outcomes data have excess zeroes. So, the spearman's correlation is also suspicious.

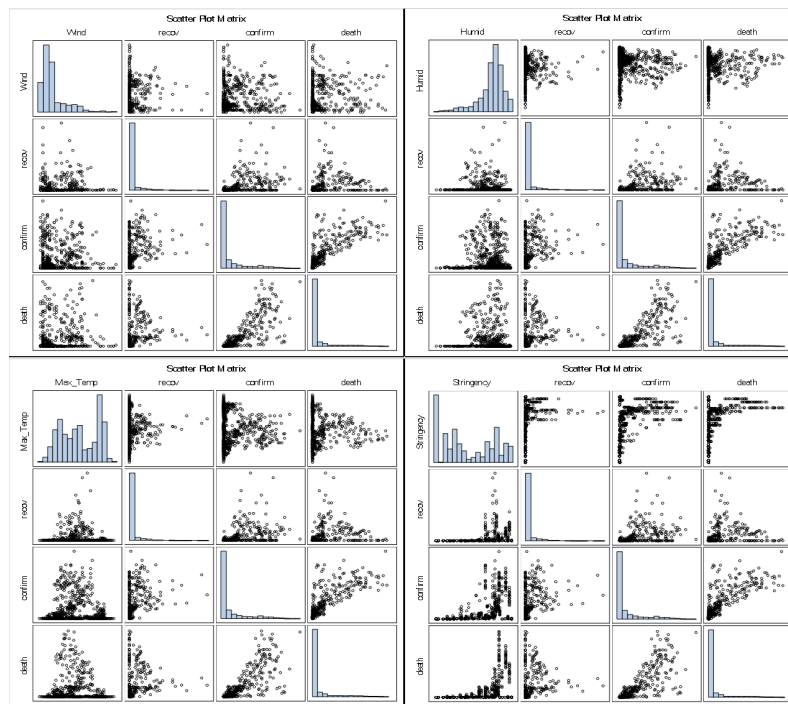
The above correlational analysis also neglects the temporal spatial nature of the data and the order in which they occur. Time is an important aspect of the data generation process. Assuming a 14 days lag, we explore the autocorrelation as well as cointegration of the spatio-temporal relationship amongst these variables. In order to demonstrate this, we generate series plot of dynamic relationship between each of the COVID-19 outcome data (see Figure 2).

One would expect these variables to be highly correlated and this is shown in Figure 2. The dynamics in time are closely related, which confirm the correlation results; recovery and confirmed cases 0.59 (0.54, 0.63) *P*<.0001, recovery and death cases 0.50 (0.46, 0.55) *P*<.0001, and confirm

**Table 2: Summary statistics of all the variables**

Variable	N	Mean	Median	Std Dev	Std Error	Minimum	Maximum
Recovery	1000	268.46	0	935.72	29.59	0	10980.00
Confirmed	1000	732.51	15.00	1486.15	47.00	0	8681.00
Death	1000	73.49	0	191.81	6.07	0	1172.00
MinTemp	1000	5.88	5.20	14.06	0.44	-41.00	27.60
MaxTemp	1000	18.81	18.00	11.16	0.35	-5.00	40.00
Humidity	1000	73.13	76.00	14.80	0.47	6.00	96.00
Windspeed	1000	5.40	3.50	5.34	0.17	0	33.30
Hosp beds/1000	1000	3.24	2.63	1.65	0.05	2.20	8.00
Hosp occupancy	1000	81.32	80.00	5.22	0.17	71.10	91.60
Stringency	1000	41.22	32.38	33.19	1.05	0	97.14
USDPC	1000	3780.60	3996.50	1604.59	50.74	1072.00	5986.00
popkm2	1000	872.10	36.50	2343.88	74.12	2.00	7894.00

Recovery = No. of recovered patients, Confirmed = No. of confirmed COVID-19 cases, Death = No. of deaths, Popkm2 = Population density per square kilometre; Hosp beds/1000 = number of hospital beds per 1000 population; Hosp occupancy = Hospital occupancy rate in percentages, and USDPC = US Dollars Per Capita spend on health. N = the total number of observations across the ten countries, i.e. 100 days per country, hence 1000.



**Figure 1: Association between meteorological parameters and COVID-19 outcomes**

and death case 0.84 (0.83, 0.86)  $P < .0001$ . However, this type of close association over time, seen in these variables, are not reflected with weather parameters, except stringency index and wind speed (see Figure 3).

According to Figure 3, stringency and death rates are strongly associated over time. Correlation estimates for stringency is 0.69 (0.66, 0.72)  $P < .0001$ . Also, death rates are also strongly associated with wind speed (top right of Figure 3), however, the correlation estimate calculated from Spearman’s correlation analysis is 0.18 (0.12, 0.24)  $P < .0001$ . Without a visual graphical inspection of these two variables, such an association would not have been captured (as Spearman’s correlation of 0.18 is weak), thus warrants a fur-

ther analysis.

On the other hand, Figure 4 presents the cross series of recovery rates and the weather parameters. Similarly, it is evident that stringency and wind speed are associated with recovery rate over time with correlation estimates, 0.60 (0.56, 0.64)  $P < .0001$  and 0.22 (0.16, 0.28)  $P < .0001$  respectively.

We generate similar plots for the confirmed cases, see Figure 5. Confirmed cases are positively associated with wind speed and stringency index over time, 0.18 (0.12, 0.24)  $P < .0001$  and 0.80 (0.78, 0.82)  $P < .0001$ , respectively.

The autocorrelation, except wind speed, with COVID-19

**Table 3: Correlation estimates between COVID-19 outcomes and meteorological factors and other variables of interest**

Variable	With variable	N	Correlation estimate	95% confidence interval		PValue for H0: Rho=0
Recovered	MaxTemp	1000	0.11	0.052	0.17	0.0003
Recovered	Humidity	1000	-0.20	-0.26	-0.14	<.0001
Recovered	Wind	1000	0.22	0.16	0.28	<.0001
Recovered	Stringency	1000	0.60	0.56	0.64	<.0001
Recovered	Hospital beds	1000	0.20	0.14	0.26	<.0001
Recovered	Hospital occupancy	1000	0.01	-0.06	0.07	0.8364
Recovered	USDPC	1000	0.06	-0.00	0.12	0.0583
Recovered	popkm2	1000	0.06	-0.01	0.12	0.0744
Confirmed	MaxTemp	1000	-0.06	-0.12	0.01	0.0755
Confirmed	Humidity	1000	-0.18	-0.24	-0.12	<.0001
Confirmed	Wind	1000	0.18	0.12	0.24	<.0001
Confirmed	Stringency	1000	0.80	0.78	0.82	<.0001
Confirmed	Hospital beds	1000	0.05	-0.01	0.11	0.1103
Confirmed	Hospital occupancy	1000	-0.06	-0.13	-0.00	0.0432
Confirmed	USDPC	1000	0.09	0.03	0.15	0.0030
Confirmed	popkm2	1000	0.17	0.11	0.23	<.0001
Death	MaxTemp	1000	-0.13	-0.19	-0.07	<.0001
Death	Humidity	1000	-0.19	-0.25	-0.13	<.0001
Death	Wind	1000	0.18	0.12	0.24	<.0001
Death	Stringency	1000	0.69	0.66	0.72	<.0001
Death	Hospital beds	1000	0.06	-0.01	0.12	0.0778
Death	Hospital occupancy	1000	-0.14	-0.20	-0.08	<.0001
Death	USDPC	1000	0.11	0.05	0.17	0.0004
Death	popkm2	1000	0.09	0.03	0.16	0.0028

Recovery = No. of recovered patients, Confirmed = No. of confirmed COVID-19 cases, Death = No. of deaths, Poppkm2 = Population density per square kilometre; Hosp beds/1000 = number of hospital beds per 1000 population; Hosp occupancy = Hospital occupancy rate in percentages, and USDPC = US Dollars Per Capita spend on health. N = the total number of observations across the ten countries, i.e. 100 days per country, hence 1000.

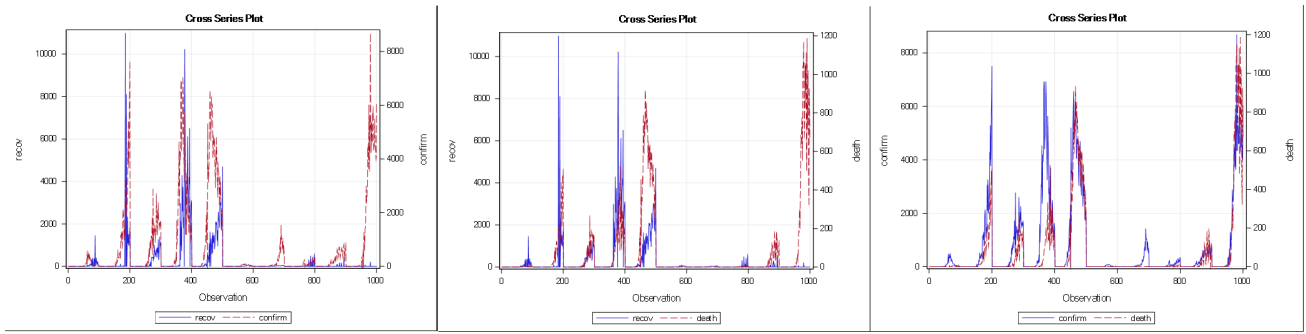


Figure 2: Spatio-temporal relationship amongst the outcomes

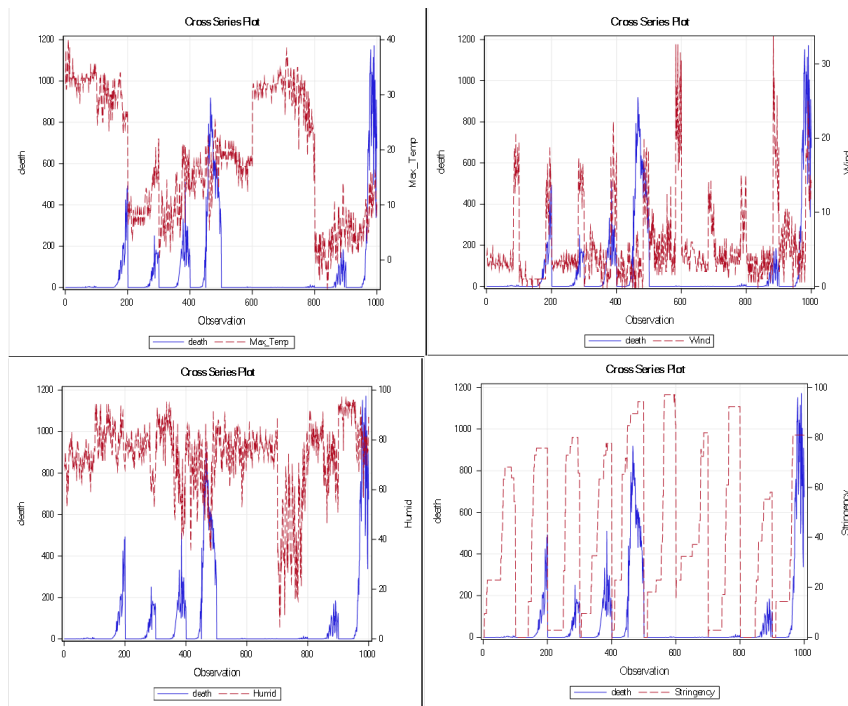


Figure 3: Spatio-temporal relationship amongst COVID-19 outcomes and meteorological factors

outcomes are random, which is ascertained by the near zero autocorrelations.

Finally, we explore relationship and causality, if any, among all the considered variables via a vector autoregressive analysis. Two groups of the variables were created, group one with recovery rates, confirmed cases and death, and group two with maximum temperature, humidity, wind speed and stringency index. The null hypothesis of the Granger causality test<sup>22</sup> is that group 1 is influenced only by itself, and not by group 2 and a p-value below 0.05 confirms the dependence.

The findings suggest that there is no evidence of dependencies among the variables with a p-value of 0.1593, which is greater than 0.05. The following relationships were established with coefficients being those with p-values below 0.10.

$$\mathbf{Recovery}(t) = 0.70 * recovery(t - 1) + 0.27 * confirm(t - 2) \quad (1)$$

$$\mathbf{Confirm}(t) = 1366.07 + 0.68 * confirm(t - 1) - 22.26 * stringency(t - 2) \quad (2)$$

$$\mathbf{Death}(t) = death(t - 1) + 2.64 * stringency(t - 1) + 0.22 * death(t - 2) - 2.84 * stringency(t - 2) \quad (3)$$

The above suggests that recovery at the current time is driven by recovery at the immediate time past and confirmed cases 2 days earlier. Also, confirmed cases at the current time is dependent on an average reported confirmed cases, plus a percentage of the immediate past confirmed cases, less a multiplicative factor of the stringency in the past two days. This seems to confirm what has been found in an unrelated work, suggesting the lock-down date only delay by 3 days, the estimated confirmed cases.<sup>23</sup> Similarly, the death reported at the current time is related to the death reported during the immediate time past, a multiplicative of death reported at the immediate time past and a multiplicative of

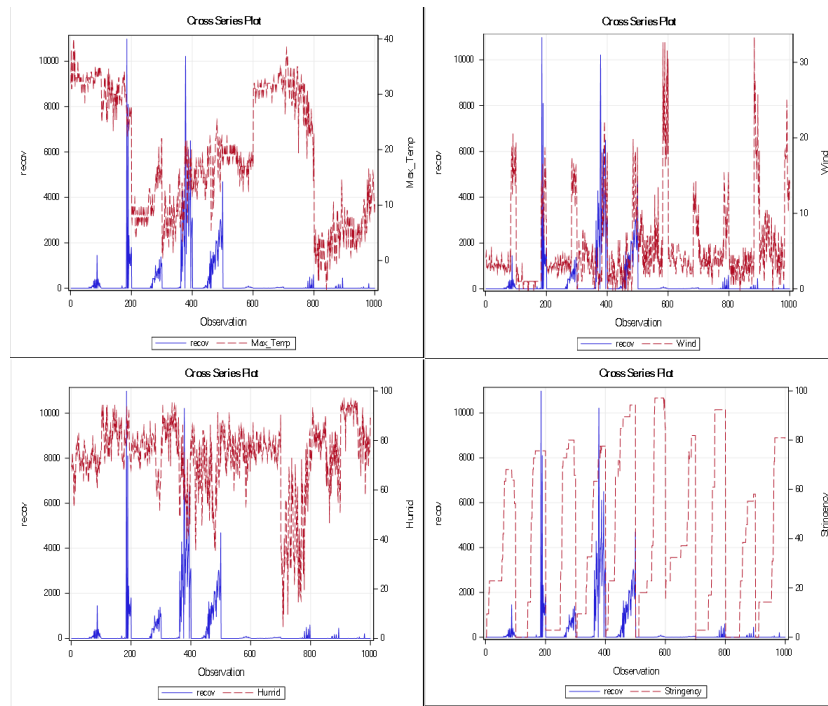


Figure 4: Cross series of recovery rates and the weather parameters and stringency index

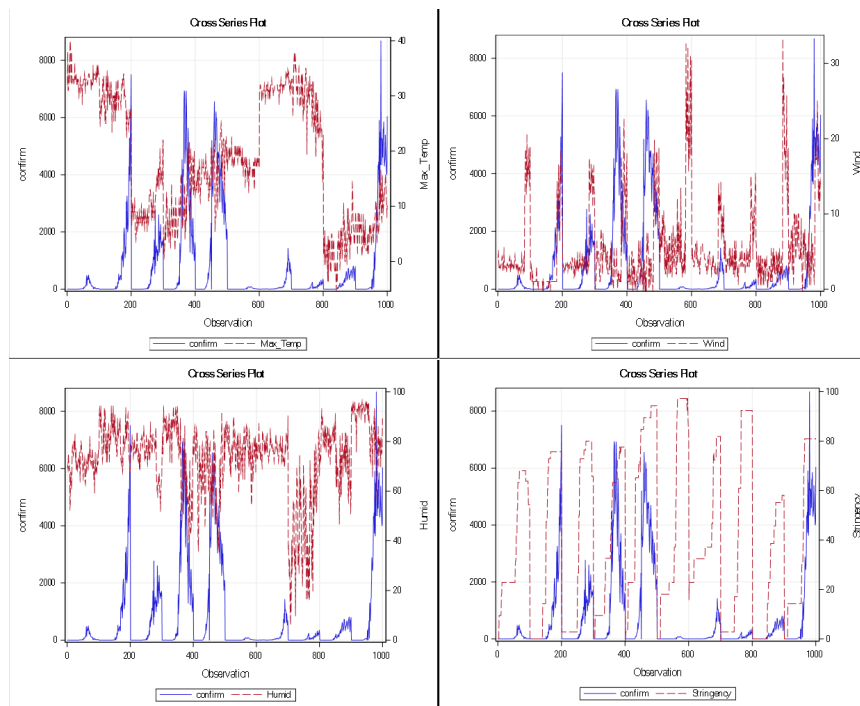


Figure 5: Cross series of daily confirmed cases and the weather parameters and stringency index

the stringency of the past two days. These results show that confirmed, recovery and death rates do not significantly depend on any of the weather parameters we considered in the research.

## DISCUSSION

The arrival of warm weather raised the expectation that the

spread of the disease will slow down and the “notorious” curve will flatten. But is it likely? Most studies in the literature noted that temperature and high levels of humidity are associated with lower negative COVID-19 incidences. However, there are serious limitations around the methodological assumptions and the data used in these studies, as explained in the introduction, thus the results can be deemed to be inaccurate and inconclusive. We therefore carried out an exhaustive analysis of data tackling some of these issues



for 10 countries from all over the world by considering all continents but Antarctica. Daily confirmed cases, number of deaths, number of recovered cases, lockdown stringency index, several meteorological factors along with other variables (number of hospital beds per 1000 population, hospital occupancy rate, GDP per capita and population density) were investigated in greater detail.

According to the graphical illustrations, there is a clear correlated pattern, between meteorological parameters and COVID-19 pandemic. Explanatory analysis of statistical results tells a different story. The findings suggest that most of the associations reported in the literature need to be interpreted with caution, as most of these articles neglected the temporal spatial nature of the data.

Based on our findings, most of the correlations no matter which coefficient is used are mostly and strictly between -0.5 and 0.5, which are weak correlations. Most of these articles reported p-values but without confidence intervals and without explaining hypotheses that have been tested. Using the Pearson correlation coefficient for our ten-country data, the data suggests weak correlation after we partial out stringency index and minimum temperature.

The Pearson correlation estimates between COVID-19 outcomes and meteorological factors suggest very weak association, ranging from -0.26 to 0.28, with wind being showing the strongest association, albeit very weak.

The correlation estimates show that stringency and death rates are strongly associated over time and similarly stringency and wind speed are associated with recovery rate. Although the correlation estimates are weak in some variables (e.g. death rates vs. wind speed), a visual inspection of the plots reveals a strong association between the variables. The strong correlation estimates for stringency index and COVID-19 outcomes could be explained as, the more cases are confirmed, government put stricter restrictions on movement of people. Also, the more confirmed cases, the more death and recovery outcomes. It suggests that confirmed outcome is a confounding factor between stringency and each of recovery and death outcome.

An interesting but not surprising finding is the correlation between stringency and each of the COVID-19 outcomes, the strongest being between stringency and confirmed cases, 0.80 (0.78, 0.82)  $P < .0001$ . However, a negative correlation would be expected between stringency and COVID-19 outcomes, as the higher SI index means a tighter lockdown. Note that countries generally increase the control (SI) as the COVID-19 cases increase. The increase in the figures generally continues for a long period of time (weeks), as a result positive correlation may arise from this trend. Lockdown policies are generally for controlling the increase in the cases. When time passes with lockdown, the figures gradually go down.

Note that there are significant variations on the national differences in COVID-19 diagnosis and reporting of outcomes (e.g. death, recovery, confirmed cases). For instance, the UK reports death in hospital and the community separately causing delay in the number of daily deaths. Also, the number of recovered patients is not available in the UK or it is reported in batches as oppose to daily in Brazil. We attempted to capture the lag and lack of the data in the reporting within our model.

Lastly, a vector autoregressive analysis is carried out to explore any relationship and causality among the variables. The analysis revealed that recovery and death rates do not significantly depend on any of the weather parameters considered in the study. In brief, the studies in the literature mainly lack evidence and clear interpretation of the statis-

tical output (e.g. underlying assumption, confidence intervals, a clear hypothesis). To uncover the devil hidden in the detail, a comprehensive and exhaustive analysis along with proper interpretation of graphical outputs and statistical analysis are essential. Based on this explorative analyses, statistical results in many articles reporting association between meteorological parameters and COVID-19 outcomes need to be interpreted with caution, especially when the spatial-temporal nature of the data generation process is ignored, the devil might be in the details.

## FUTURE WORKS

As an extension of this study, we will jointly model longitudinal meteorological factors with COVID-19 outcomes, adjusted for key covariates, such as GDP per capita, number of intensive care unit (ICU) beds, deprivation score and stringency index. Thus, it is expected that (1) the trajectory over time (evolution) of COVID-19 incidence will be related to the evolution of a meteorological factor, as one may have direct and indirect implications on the other, and (2) the effect of covariates may not be captured if we ignore the interdependence in the evolution of these outcomes. The advantages of the joint modelling approach are the ability of accounting for interdependence by bringing together the models for each outcome (both COVID-19 incidences and meteorological factor) by specifying a joint distribution for their random effect terms.

## CONCLUSION

Our exhaustive analysis showed that the associations reported in the literature need to be interpreted with caution, especially when the spatial-temporal nature of the data generation process is ignored, the devil might be in the details. A notable shortcoming of this study is the period for which the data have been collected and the variability in the countries considered, especially in the position along the course of the pandemic. The authors think that, time will tell if really there are any strong associations between meteorological parameters and the pandemic. However, without effective control measures, strong outbreaks are likely in more windy climates and summer weather, humidity or warmer temperature will not substantially limit pandemic growth. Therefore, we postulate that the drivers would be more related to the way of life, urban population density, healthcare system preparedness and socio-economics characteristics.

**Funding:** None

### Authorship contributions:

All authors have contributed to the study equally. The manuscript has been reviewed, edited, and approved by all authors.

### Competing interests:

The authors completed the Unified Competing Interest form at [www.icmje.org/coi\\_disclosure.pdf](http://www.icmje.org/coi_disclosure.pdf) (available upon request from the corresponding author), and declare no conflicts of interest.

**Correspondence to:**

Usame Yakutcan  
PhD, AFHEA, MSc, BSc  
Hertfordshire Business School  
University of Hertfordshire

Hatfield, UK  
[u.yakutcan@herts.ac.uk](mailto:u.yakutcan@herts.ac.uk)



This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (CC-BY-4.0). View this license's legal deed at <http://creativecommons.org/licenses/by/4.0> and legal code at <http://creativecommons.org/licenses/by/4.0/legalcode> for more information.

## REFERENCES

1. Ghalhari GF, Mayvaneh F. Effect of Air Temperature and Universal Thermal Climate Index on Respiratory Diseases Mortality in Mashhad, Iran. *Archives of Iranian Medicine (AIM)*. 2016;19(9).
2. Dadbakhsh M, Khanjani N, Bahrampour A, Haghighi PS. Death from respiratory diseases and temperature in Shiraz, Iran (2006-2011). *International journal of biometeorology*. 2017;61(2):239-246. doi:10.1007/s00484-016-1206-z
3. Gomez-Acebo I, Llorca J, Dierssen T. Cold-related mortality due to cardiovascular diseases, respiratory diseases and cancer: A case-crossover study. *Public Health*. 2013;127(3):252-258. doi:10.1016/j.puhe.2012.12.014
4. Islam N, Shabnam S, Erzurumluoglu AM. Temperature, humidity, and wind speed are associated with lower Covid-19 incidence. *medRxiv*. March 2020. doi:10.1101/2020.03.27.20045658
5. Ma Y, Zhao Y, Liu J, He X, Wang B, Fu S, et al. Effects of temperature variation and humidity on the death of COVID-19 in Wuhan, China. *Science of The Total Environment*. 2020:138226. doi:10.1016/j.scitotenv.2020.138226
6. Chen B, Liang H, Yuan X, et al. Roles of meteorological conditions in COVID-19 transmission on a worldwide scale. *MedRxiv*. 2020. doi:10.1101/2020.03.16.20037168
7. Oliveiros B, Caramelo L, Ferreira NC, Caramelo F. Role of temperature and humidity in the modulation of the doubling time of COVID-19 cases. *medRxiv*. 2020. doi:10.1101/2020.03.05.20031872
8. Notari A. Temperature dependence of COVID-19 transmission. *medRxiv*. 2020. doi:10.1101/2020.03.26.20044529
9. Sajadi MM, Habibzadeh P, Vintzileos A, Shokouhi S, Miralles-Wilhelm F, Amoroso A. Temperature and latitude analysis to predict potential spread and seasonality for COVID-19. *SSRN*. 2020. doi:10.2139/ssrn.3550308
10. Wang J, Tang K, Feng K, Lv W. High temperature and high humidity reduce the transmission of COVID-19. *SSRN*. 2020. doi:10.2139/ssrn.3551767
11. Yao Y, Pan J, Liu Z, Meng X, Wang W, Kan H, et al. No Association of COVID-19 transmission with temperature or UV radiation in Chinese cities. *European Respiratory Journal*. 2020;55(5). doi:10.1183/13993003.00517-2020
12. Xie J, Zhu Y. Association between ambient temperature and COVID-19 infection in 122 cities from China. *Science of The Total Environment*. 2020:138201. doi:10.1016/j.scitotenv.2020.138201
13. Bannister-Tyrrell M, Meyer A, Faverjon C, Cameron A. Preliminary evidence that higher temperatures are associated with lower incidence of COVID-19, for cases reported globally up to 29th February 2020. *medRxiv*. 2020. doi:10.1101/2020.03.18.20036731
14. Bashir MF, Ma B, Komal B, Bashir MA, Tan D, Bashir M. Correlation between climate indicators and COVID-19 pandemic in New York, USA. *Science of The Total Environment*. 2020:138835. doi:10.1016/j.scitotenv.2020.138835
15. Sahin M. Impact of weather on COVID-19 pandemic in Turkey. *Science of The Total Environment*. 2020:138810. doi:10.1016/j.scitotenv.2020.138810
16. Tosepu R, Gunawan J, Effendy DS, Lestari H, Bahar H, Asfian P. Correlation between weather and Covid-19 pandemic in Jakarta, Indonesia. *Science of The Total Environment*. 2020:138436. doi:10.1016/j.scitotenv.2020.138436
17. Remuzzi A, Remuzzi G. COVID-19 and Italy: What next? *The Lancet*. 2020;395(10231):1225-1228. doi:10.1016/s0140-6736(20)30627-9
18. Hale T, Webster S, Petherick A, Phillips T, Kira B. Oxford COVID-19 Government Response Tracker, Blavatnik School of Government. Attribution CC BY standard. <https://www.bsg.ox.ac.uk/research/research-projects/coronavirus-government-response-tracker>. Published 2020.
19. OECD. *Health at a Glance*. OECD; 2019. [https://www.oecd-ilibrary.org/social-issues-migration-health/health-at-a-glance-2019\\_4dd50c09-en](https://www.oecd-ilibrary.org/social-issues-migration-health/health-at-a-glance-2019_4dd50c09-en).
20. Stenberg K, Lauer JA, Gkountouras G, Fitzpatrick C, Stanciole A. Econometric estimation of WHO-CHOICE country-specific costs for inpatient and outpatient health service delivery. *Cost Effectiveness and Resource Allocation*. 2018;16(1):11. doi:10.1186/s12962-018-0095-x
21. Liu Q, Li C, Wanga V, Shepherd BE. Covariate - adjusted Spearman's rank correlation with probability - scale residuals. *Biometrics*. 2018;74(2):595-605. doi:10.1111/biom.12812

22. Engle RF, Granger CW. Co-integration and error correction: Representation, estimation, and testing. *Econometrica: journal of the Econometric Society*. 1987;251-276. [doi:10.2307/1913236](https://doi.org/10.2307/1913236)

23. Yuan C. A simple model to assess Wuhan lockdown effect and region efforts during COVID-19 epidemic in China Mainland [Submitted]. *Bull World Health Organ*. 2020.

## FIGURES, TABLES, AND SUPPLEMENTARY MATERIALS

### Online Supplementary Document

Download: <https://www.joghr.org/article/13693-a-statistical-assessment-of-association-between-meteorological-parameters-and-covid-19-pandemic-in-10-countries/attachment/37880.docx>

---