

1 **Title: INDEPENDENT ASSOCIATION OF METEOROLOGICAL CHARACTERISTICS**  
2 **WITH INITIAL SPREAD OF COVID-19 IN INDIA**

3 **Authors:**

4 **Hemant Kulkarni, MBBS, MD<sup>1,2,\*</sup>; Harshwardhan V. Khandait, MBBS<sup>3,\*</sup>; Uday W.**  
5 **Narlawar, MBBS, MD<sup>1,3</sup>; Pragati G. Rathod, MBBS, MD<sup>3</sup>; Manju Mamtani, MBBS, MD<sup>1,2</sup>**

6

7 **Affiliations:**

8 <sup>1</sup>Lata Medical Research Foundation, Nagpur, India

9 <sup>2</sup>M&H Research, LLC, San Antonio, Texas, USA

10 <sup>3</sup>Government Medical College, Nagpur, India

11 \*, Equal contribution as first author

12 **Article Summary Line:** Although air temperature and wind speed were significantly associated  
13 with initial COVID-19 transmission in India, the contribution of lockdown to variability of the  
14 basic reproduction rate was stronger.

15 **Running Title:** Meteorological factors and COVID-19 in India

16 **Keywords:** temperature, wind speed, COVID-19, basic reproduction rate

17

18 **Address for correspondence:** Hemant Kulkarni, MD; 12023 Waterway Rdg, San Antonio,  
19 Texas, USA, 78249. email: [hemant.kulkarni@mhresearch.com](mailto:hemant.kulkarni@mhresearch.com).

20

21 **Abstract—word count (133)**

22           Whether weather plays a part in the transmissibility of the novel COronaVirus Disease-  
23 19 (COVID-19) is still not established. We tested the hypothesis that meteorological factors (air  
24 temperature, relative humidity, air pressure, wind speed and rainfall) are independently  
25 associated with transmissibility of COVID-19 quantified using the basic reproduction rate ( $R_0$ ).  
26 We used publicly available datasets on daily COVID-19 case counts (total  $n = 108,308$ ), three-  
27 hourly meteorological data and community mobility data over a three-month period. Estimated  
28  $R_0$  varied between 1.15-1.28. Mean daily air temperature (inversely) and wind speed (positively)  
29 were significantly associated with time dependent  $R_0$ , but the contribution of countrywide  
30 lockdown to variability in  $R_0$  was over three times stronger as compared to that of temperature  
31 and wind speed combined. Thus, abating temperatures and easing lockdown may concur with  
32 increased transmissibility of COVID-19.

33

34 **Text—word count (3216)**

## 35 **INTRODUCTION**

36 As the novel CORonaVirus Disease-19 (COVID-19) continues to devastate the world,  
37 there remains a myriad of unknowns about its pathogenesis, population dynamics, epidemiology,  
38 prevention and treatment. Since its introduction into the global susceptible population SARS-  
39 CoV-2, the causative agent of COVID-19, has presented several conundrums. It was initially  
40 believed that like many other viruses, SARS-CoV-2 may also be responsive to the environmental  
41 influences posed by climatic and meteorological factors<sup>1-4</sup>. However, current understanding of  
42 the potential role of weather on the spread of SARS-CoV-2 is far from clear.

43 The COVID-19 outbreaks have been generally more severe in the countries located in the  
44 mid-latitudes where the temperature is considerably low in contrast to the tropical countries.  
45 Several studies around the world have attempted to specifically establish a relationship between  
46 COVID-19 transmission and various meteorological factors.<sup>5-7</sup> For example, a study conducted  
47 in New York, USA, found that mean temperature, minimum temperature and air quality had a  
48 significant association with the COVID-19 pandemic.<sup>5</sup> Similarly, Shi et al<sup>8</sup> reported a  
49 statistically significant correlation between daily temperature and daily count of COVID-19  
50 cases in China and suggested that temperatures above 8–10 °C would lead to a decline in the  
51 number of infected cases. In a parallel investigation, Prata et al<sup>5</sup> concluded that a rise in 1 °C  
52 temperature would result in a decrease in the number of daily confirmed COVID-19 cases in  
53 Brazil. There have been very few investigations from India in this regard<sup>9-11</sup> – a country with  
54 second largest population size after China. These studies from India have generally indicated a  
55 potential role of weather conditions in the spread of COVID-19.

56           On the other end of the spectrum, a study conducted by Yao et al<sup>12</sup> concluded that there is  
57 no association of COVID-19 transmission with temperature or UV radiation in Chinese cities.  
58 Indeed, an elegant, evidence-based review by Brassley et al<sup>6</sup> summarized the existing evidence  
59 in this regard and observed that a. cold and dry conditions may facilitate the spread of the novel  
60 coronavirus (2019-nCoV) b. much of the emerging data for 2019-nCoV has yet to be peer-  
61 reviewed and is thus needed; and c. relying on weather changes alone to slow the transmission of  
62 COVID-19 are unlikely to be sufficient. Considering these recommendations; the variability in  
63 the observed associations; and a relative lack of such studies from India, we conducted this  
64 investigation on a nationwide sample of geographical locations across India. The primary goal  
65 was to test the putative association of geo-meteorological characteristics with rates of COVID-19  
66 transmission and to test its independence from other socio-behavioral interventions like  
67 lockdowns and mobility.

## 68 **MATERIALS AND METHODS**

69           **Data sources:** We selected a total of 46 geographical locations across India. For each  
70 selected location (either a city, union territory or district), we collected data for a three-month  
71 period (March 1, 2020 through May 31, 2020). Following data items were collected for each  
72 study location: daily number of confirmed COVID-19 cases, meteorological data, demographic  
73 data and overall geographic data. The meteorological data included 3-hourly recordings of  
74 temperature, relative humidity, air pressure, wind speed and rainfall. The demographic data  
75 included the 2011 census population and the geographical data included area and elevation. The  
76 area and population records were combined to estimate the community density. Lastly, temporal  
77 data on the lockdown implementation phases and the mobility of the population (estimated

78 anonymously from the cellphone use data) was collected to study the potential temporal  
79 concurrence with COVID-19 transmission.

80 All data used in this study are publicly available and are completely anonymized. The  
81 study was approved by the Institutional Ethics Committee of Government Medical College,  
82 Nagpur, India. Following were the sources of data: number of daily COVID-19 cases –  
83 <https://api.covid19india.org/>; meteorological data –<https://www.tutiempo.net/> and  
84 <https://www.worldweatheronline.com/>; 2011 census data – [https://censusindia.gov.in/2011-](https://censusindia.gov.in/2011-common/censusdata2011.html)  
85 [common/censusdata2011.html](https://censusindia.gov.in/2011-common/censusdata2011.html); and geographical data – combination of census data and search  
86 on Wikipedia® (<https://en.wikipedia.org/wiki/Wikipedia>). Lastly, the temporal mobility data was  
87 downloaded from the publicly available repository: <https://www.google.com/covid19/mobility/>.  
88 These indicated percent change from baseline mobility on visits to the following five  
89 destinations - retail and recreation, grocery and pharmacy, parks, transit stations and workplaces.

90 **Quantification of COVID-19 transmissibility:** Using the daily case count data we  
91 estimated the basic reproduction rate ( $R_0$ ) in two different ways. First, we estimated the average  
92  $R_0$  over the entire duration of 92 days period of data collection. For this, we used two methods –  
93 the exponential growth (EG) and the maximum likelihood (ML). Second, we estimated the daily  
94  $R_0$  in a time-dependent (TD) fashion. All estimates of  $R_0$  require a knowledge of serial interval,  
95 the time difference between onset of symptoms in an infector and an infectee. We assumed a  
96 gamma distributed serial interval with a mean of 3.96 days and a standard deviation of 4.75 days  
97 as reported by Du et al.<sup>13</sup> We used the R package  $R_0$ <sup>14</sup> to derive all the estimates of  $R_0$ . Finally,  
98 we considered the possibility of biased estimates of  $R_0$  owing to the relative lack of testing  
99 facilities, especially during the initial period of the epidemic. For this, we used the method of

100 Lachmann et al<sup>15</sup> that considers South Korea as the reference country and estimates the degree of  
101 undertesting by combining demographic and vital statistics data. Using this method, we derived  
102 the possible undertesting on each study day.

103 **Statistical analysis:** Our analyses used estimates of  $R_0$  as the dependent variable and the  
104 geo-meteorological and socio-behavioral characteristics as the explanatory variables. To  
105 compare groupwise means we used the Mann-Whitney U test or Kruskal-Wallis test as  
106 appropriate. Significance of heterogeneity across study locations was statistically tested using the  
107 Q test. Time series data were smoothed using a five-day sliding window technique. Further, to  
108 make the different time series (each meteorological characteristic) comparable, we converted  
109 them to a series of z-scores. To test the temporal concurrence, we used the cross-correlation  
110 between two time series (Pearson's correlation). To test the association of time series variables  
111 with estimated time dependent  $R_0$ , we used multivariable, ordinary least squares regression.  
112 Starting with the full model, we conducted stepwise, backward elimination regression modeling  
113 with a probability retention criterion of 0.05. Lastly, to quantify the relative contribution of each  
114 covariate with time dependent  $R_0$ , we estimated the proportional reduction in error (PRE) using  
115 the approach of Judd, McClelland and Ryan.<sup>16</sup> PRE was estimated as reduction in the residual  
116 sum of squares by including a covariate in the full model. Statistical analyses were conducted  
117 using the Stata 14.2 statistical package (Stata Corp, College Station, TX). Type 1 error rate of  
118 0.05 was used for hypothesis testing.

## 119 **RESULTS**

120 **Representativeness of the study locations:** We included 46 locations across India that  
121 contained 32 cities, 12 districts and 2 union territories. Figure 1 shows the geographical spread of

122 these locations and the geographic and demographic details for these locations are provided in  
123 Supplementary Table 1. The study locations varied widely in terms of the area (range 4.86 –  
124 6039 sq. miles), elevation (range 3 – 11500 feet above sea level) and population density (range  
125 33.6 to 56812.3/sq. mile). The selected locations are distributed across India and represent  
126 majority of the states / union territories of India. Meteorological data was available on all the  
127 selected study locations.

128         The cumulative number of COVID-19 confirmed cases (till and including May 31, 2020)  
129 reported from these locations also varied widely (1 to 37666). The 46 selected locations together  
130 accounted for a total of 108,308 confirmed COVID-19 cases. From entire India the number of  
131 cumulative COVID-19 cases till May 31, 2020 were 182,140. Thus, our selected geographic  
132 locations accounted for ~60% of all India COVID-19 cases till May 31, 2020. The top 5  
133 contributing locations to the overall cumulative COVID-19 case counts were Mumbai (37666),  
134 Delhi (18058), Chennai (12040), Ahmedabad (11919) and Pune (7459) as shown in  
135 Supplementary Figure 1.

136         **Average estimated  $R_0$  for COVID-19:** We first estimated the  $R_0$  based on case counts  
137 reported for the entire country as well as only for the locations included in this study. For each of  
138 these datasets, we estimated the  $R_0$  in two ways – first based on the actual reported case counts  
139 and second by inflating the case counts to account for the potential undertesting on each day. The  
140 results of these analyses are shown in Figure 2 and referred to as unadjusted (actual case counts,  
141 blue bars) and adjusted (for potential undertesting, purple bars). Our average estimates of  $R_0$   
142 using different methods of estimation and with or without adjusting for undertesting ranged from  
143 1.18 to 1.27 for India and 1.15 to 1.28 for the selected study locations. All the estimates and their

144 95% confidence intervals (error bars in Figure 2) were significantly above unity. These results  
145 indicated that over the study period, the average estimates of  $R_0$  were significantly greater than  
146 one, confirming the existence of the epidemic; the average  $R_0$  estimates were only moderately  
147 above unity; the average  $R_0$  estimates were minimally influenced by potential undertesting; and  
148 that the study locations yielded average  $R_0$  estimates consistent with those for the whole country  
149 thereby indirectly reaffirming the representativeness of the selected study locations.

150 We also examined the heterogeneity of the average  $R_0$  estimates across the study  
151 locations. For these analyses, we restricted the locations which showed at least seven consecutive  
152 days with a contiguous segment of non-zero cases. Total of 35 locations were eligible based on  
153 this criterion. The average  $R_0$  estimates derived using the ML method [point estimates and  
154 confidence intervals (CI)] for these 35 locations are shown in Figure 3. There was a significant  
155 heterogeneity in the average  $R_0$  estimates ( $p = 6.9 \times 10^{-30}$ ) with estimates ranging from 1.98 for  
156 Dehradun to 0.89 for Kolkata. The average  $R_0$  estimates for the top five contributing locations  
157 were: Mumbai 1.16 (95% CI 1.14 – 1.18); Delhi 1.25 (95% CI 1.23 – 1.28); Chennai 1.20 (95%  
158 CI 1.17 – 1.23); Ahmedabad 1.10 (95% CI 1.07 – 1.13) and Pune 1.22 (95% CI 1.18 – 1.26).

159 **Temporal changes in  $R_0$  estimates:** Next, we considered the variability in  $R_0$  estimates  
160 over the duration of the study for all locations together. Figure 4 shows that the  $R_0$  estimates  
161 were initially high but undulated widely and gradually converged towards the overall estimates  
162 shown in Figure 2 with narrow confidence bands later. Thus, the time dependent  $R_0$  estimates  
163 showed considerable variation across study time.

164 We examined the association of the time-dependent  $R_0$  estimates with two socio-  
165 behavioral characteristics – implementation of a countrywide lockdown and the extent of social

166 distancing as reflected by the cellphone mobility data. When contrasted against the various  
167 phases of countrywide lockdown in India (grey shaded regions in Figure 4), we found that the  
168 median  $R_0$  estimates consistently reduced as lockdown was imposed. Before lockdown began  
169 (March 1 through March 24, 2020) the median  $R_0$  estimate was 1.54 and this estimate decreased  
170 to 1.40 (March 25 – April 14, 2020), 1.21 (April 15 – May 3, 2020), 1.16 (May 4 – May 17,  
171 2020) and 1.10 (May 18, 2020 onwards) during lockdown phases 1 through 4, respectively  
172 (Kruskal-Wallis  $p < 0.0001$ ).

173 The cellphone-based community mobility data also revealed consistent and interesting  
174 patterns. As shown in Supplementary Figure 2, the overall trends in community mobility for all  
175 five destinations showed a dramatic decrease around the beginning of phase 1 lockdown,  
176 remained very low during phase 1 lockdown and then gradually increased as the lockdown  
177 progressed. The 5-day rolling z-scores for the average mobility based on these five parameters is  
178 shown in Figure 4 (green curve).

179 **Association of time dependent estimates with meteorological data:** The time trends  
180 for temperature, relative humidity, air pressure, wind speed and rainfall are shown in Figure  
181 5A. Over the duration of the study, temperature and wind speed steadily increased; relative  
182 humidity and air pressure gradually decreased while rainfall remained steady. As a first step of  
183 the association analyses, we estimated the cross-correlation between each meteorological  
184 variable and the  $R_0$  estimates. Figure 5B shows the cross-correlograms for lags ranging from -10  
185 to 10 days. We found that higher temperature, wind speed and rainfall were correlated inversely  
186 while relative humidity and air pressure were correlated positively with time dependent  $R_0$   
187 estimates. The best cross-correlation was observed for temperature and humidity on the same day

188 (-0.73 and 0.63, respectively), wind speed on previous day (-0.40), rainfall preceding by 4 days  
189 (-0.29) and air pressure preceding by 6 days (0.54). Together these results indicated that  
190 concurrent or immediately preceding values of meteorological variables are significantly  
191 correlated with time dependent  $R_0$  estimates.

## 192 **Multivariable association of meteorological and socio-behavioral predictors with**

193 **time dependent  $R_0$ :** We then examined whether the meteorological and socio-behavioral

194 covariates were independently associated with time dependent  $R_0$  estimates. The full regression

195 model used time dependent  $R_0$  estimates as the dependent variable and following 14 covariates

196 as explanatory variables: five z-scores for the meteorological covariates, five z-scores for

197 community mobility data and four phases of lockdown (each used as a dichotomous variable).

198 The results of these analyses are shown in Table 1. In the full model, we observed that the

199 lockdown phases 3 (only marginally) and 4 and wind speed were the only covariates that were

200 statistically significantly associated with  $R_0$  estimates. In this context, the mobility data (which

201 was highly correlated with the lockdown phases) did not retain statistical significance. However,

202 considering the potential for interactions among covariates and the possibility of an

203 underpowered full model (14 covariates observed on 92 days), we conducted stepwise regression

204 modeling with a probability retention criterion of 0.05. The results of the final model (Table 1)

205 showed that temperature z-scores, wind speed and lockdown phases 2-4 were retained in the

206 final model. This model fitted the data well with an adjusted  $R^2$  of 0.56 (Supplementary Figure

207 3)

208 From the point of public health relevance, we then quantified the contribution of each

209 variable retained in the final model to the overall variance of time dependent  $R_0$ . The PRE

210 estimates for the variables retained in the final model were as follows: temperature: 9.1%, wind  
211 speed: 9.9%, lockdown phase 2: 12.2%, lockdown phase 3: 22.5% and lockdown phase 4:  
212 27.0%. These results indicate that while the meteorological factors of temperature and wind  
213 speed were statistically significant predictors of COVID-19 transmissibility, their contribution to  
214 dampening the  $R_0$  estimate was 3-4 times weaker as compared to the countrywide lockdown  
215 phases 2-4.

## 216 **DISCUSSION**

217 Using nationwide data from India over a three-month period, our study made three  
218 cardinal observations. First, the average basic reproduction rate ( $R_0$ ) of COVID-19 infection in  
219 the period from March 1 through May 31, 2020 ranged from 1.15 to 1.28 even after accounting  
220 for the potential undertesting. Second, the COVID-19 transmissibility (quantified using  $R_0$ ) was  
221 significantly associated with daily average temperature (inversely), daily average wind speed  
222 (positively) and the countrywide intervention of lockdown (inversely). Third, the contribution of  
223 lockdown to the variability in time dependent  $R_0$  was three times more than the contribution of  
224 temperature and wind speed combined. Together, these results suggest that in India while the  
225 meteorological determinants of COVID-19 were independently associated with the  
226 transmissibility, their contribution was outweighed by that of the countrywide lockdown.

227 Even though statistically significantly greater than unity, our estimate of  $R_0$  was low.  
228 This estimate is comparable to the value of 1.32 reported by Du et al. However, the low value of  
229  $R_0$  should be interpreted with caution. First, there has been a debate about the length of serial  
230 interval with values ranging from as low as 3 days to as high as 9 days.<sup>13, 17-21</sup> We used the serial  
231 interval of ~4 days which is on the lower side of the serial interval range and could have partly

232 contributed to the low  $R_0$  observed in this study. Second, the major part of the study period  
233 included lockdown and reduced mobility and therefore the  $R_0$  estimate may represent a muted  
234 transmissibility owing to interventions in place. Third, the low  $R_0$  estimate does not indicate lack  
235 of viral infectiousness or any other viral characteristic but only implies the extent of potential  
236 spread of the disease.<sup>22</sup> Fourth, the epidemic of COVID-19 is still ongoing and our estimate of  $R_0$   
237 only captures the initial, ascending limb of the epidemic curve. Therefore, this  $R_0$  estimate does  
238 not fully capture the population dynamics of COVID-19. Fifth, our estimate of  $R_0$  is a  
239 conglomerate of the varying estimates across the study locations as shown in Figure 3. The  
240 variability in  $R_0$  across study locations indicates that the location-specific epidemic curves were  
241 not aligned to the same starting point in time and therefore our  $R_0$  estimate should not be used as  
242 a generalizable estimate of COVID-19 transmissibility. The reason for estimating  $R_0$  in the study  
243 was to investigate the potential influence of geo-meteorological factors on transmissibility.

244         Several researchers around the world have demonstrated an inverse relationship between  
245 air temperature and COVID-19 transmissibility.<sup>23-27</sup> Our results also are in agreement with the  
246 general understanding that higher ambient temperature can inversely influence COVID-19  
247 transmissibility.<sup>23, 27</sup> The study duration mark a period of increasing temperature in the Indian  
248 peninsula and our results indicate that, in general, high ambient temperatures were associated  
249 with lower  $R_0$  estimates such that unit standard deviation increase in air temperature was  
250 associated with a 0.08 lower  $R_0$  (Table 1, final model). On the other hand, we observed that a unit  
251 standard deviation increase in wind speed was associated with a 0.08 higher  $R_0$  (Table 1, final  
252 model). The current evidence for the potential role of wind speed in COVID-19 spread is  
253 conflicting with studies reporting positive,<sup>28</sup> null<sup>29-31</sup> and negative<sup>1, 32</sup> association with COVID-  
254 19 transmissibility. Our observation of a positive association of COVID-19 transmissibility with

255 wind speed is in line with the growing idea that the SARS-CoV-2 virus may be airborne.<sup>33, 34</sup> Of  
256 note, incidence of COVID-19 has been shown to be associated with air pollution<sup>1, 35, 36</sup> – a factor  
257 that is significantly influenced by wind speed.<sup>37</sup> Our study cannot directly answer these  
258 interesting hypotheses, which should be tested in future studies. Nevertheless, a head-to-head  
259 comparison indicated that the lockdown period was associated with three times stronger  
260 contribution to the variability in  $R_0$  as compared to that of air temperature and wind speed  
261 combined. From the perspective of public health action, this observation supports the role of  
262 proactive interventions to de-escalate the transmissibility of COVID-19. Conceivably, as the  
263 temperatures wane and the lockdown is eased, more cases of COVID-19 can be expected.

264 Our results should be interpreted in the light of some limitations. First, this was a  
265 retrospective analysis that combined data from different sources. The data are collected at the  
266 level of geographic locations and not at the level of individual patient. For example, person-to-  
267 person transmissibility of COVID-19 in an infector-infectee scenario was not investigated in this  
268 study. Therefore, all the estimates and associations should only be considered as general patterns  
269 rather than definitive evidence. Second, akin to any observational study, unmeasured  
270 confounding can be expected to be operational. Despite these potential limitations our study  
271 demonstrated interesting and important patterns of association of geo-meteorological factors in  
272 COVID-19 spread. To control a pandemic of this magnitude, all scientific evidence from a  
273 holistic standpoint is needed. To that end, our study provides clues into the ecological aspects of  
274 COVID-19 during the initial months in India.

275

276

277 **Author Bio**

278 Dr. Kulkarni is the Chief Executive Officer of M&H research, LLC, in San Antonio,  
279 Texas, USA and is also the President of the not-for-profit Lata Medical Research Foundation  
280 based in Nagpur, India. Dr. Kulkarni has published over 130 papers on various aspects of  
281 biomedical research including non-communicable diseases, HIV and AIDS, genetics,  
282 environmental sciences and mathematical modeling of infectious diseases.

283

284 **References**

- 285 1. Adhikari A, Yin J. Short-Term Effects of Ambient Ozone, PM2.5, and Meteorological  
286 Factors on COVID-19 Confirmed Cases and Deaths in Queens, New York. *International*  
287 *journal of environmental research and public health* 2020; **17**.
- 288
- 289 2. Briz-Redon A, Serrano-Aroca A. A spatio-temporal analysis for exploring the effect of  
290 temperature on COVID-19 early evolution in Spain. *The Science of the total environment*  
291 2020; **728**: 138811.
- 292
- 293 3. Qi H, Xiao S, Shi R, *et al.* COVID-19 transmission in Mainland China is associated with  
294 temperature and humidity: A time-series analysis. *The Science of the total environment*  
295 2020; **728**: 138778.
- 296
- 297 4. Tosepu R, Gunawan J, Effendy DS, *et al.* Correlation between weather and Covid-19  
298 pandemic in Jakarta, Indonesia. *The Science of the total environment* 2020; **725**: 138436.
- 299
- 300 5. Prata DN, Rodrigues W, Bermejo PH. Temperature significantly changes COVID-19  
301 transmission in (sub)tropical cities of Brazil. *The Science of the total environment* 2020;  
302 **729**: 138862.
- 303
- 304 6. Brassley J, Heneghan C, Mahtani KR, *et al.*: Do weather conditions influence the  
305 transmission of the coronavirus (SARS-CoV-2)? In, Oxford, London, Centre for

- 306 Evidence-Based Medicine, Nuffield and Department of Primary Care Health Sciences,  
307 University of Oxford, 2020  
308
- 309 7. Vantarakis A, Chatziprodromidou I, Apostolou T. COVID-19 and Environmental factors.  
310 A PRISMA-compliant systematic review. *medRxiv : the preprint server for health*  
311 *sciences* 2020: 2020.2005.2010.20069732.  
312
- 313 8. Shi P, Dong Y, Yan H, *et al.*). Impact of temperature on the dynamics of the COVID-19  
314 outbreak in China. *The Science of the total environment* 2020; **728**: 138890.  
315
- 316 9. Das K, Chatterjee ND. Examine the impact of weather and ambient air pollutant  
317 parameters on daily case of COVID-19 in India. *medRxiv : the preprint server for health*  
318 *sciences* 2020: 2020.2006.2008.20125401.  
319
- 320 10. Singh K, Agarwal A. Impact of weather indicators on the COVID-19 outbreak: A multi-  
321 state study in India. *medRxiv : the preprint server for health sciences* 2020:  
322 2020.2006.2014.20130666.  
323
- 324 11. Gupta A, Pradhan B. Impact of Daily Weather on COVID-19 outbreak in India. *medRxiv*  
325 *: the preprint server for health sciences* 2020: 2020.2006.2015.20131490.  
326
- 327 12. Yao Y, Pan J, Liu Z, *et al.* No association of COVID-19 transmission with temperature or  
328 UV radiation in Chinese cities. *The European respiratory journal* 2020; **55**.

- 329
- 330 13. Du Z, Xu X, Wu Y, *et al.* Serial Interval of COVID-19 among Publicly Reported  
331 Confirmed Cases. *Emerging infectious diseases* 2020; **26**: 1341-1343.
- 332
- 333 14. Obadia T, Haneef R, Boelle PY. The R0 package: a toolbox to estimate reproduction  
334 numbers for epidemic outbreaks. *BMC medical informatics and decision making* 2012;  
335 **12**: 147.
- 336
- 337 15. Lachmann A, Jagodnik KM, Giorgi FM, *et al.* Correcting under-reported COVID-19 case  
338 numbers: estimating the true scale of the pandemic. *medRxiv : the preprint server for*  
339 *health sciences* 2020: 2020.2003.2014.20036178.
- 340
- 341 16. Judd CM, McClelland GH, Ryan CS. *Data analysis: A model comparison approach*, 2nd  
342 edn. Routledge/Taylor & Francis Group, 2009.
- 343
- 344 17. Moradi Y, Eshrati B. Estimation of the net reproductive number of COVID-19 in Iran.  
345 *Medical journal of the Islamic Republic of Iran* 2020; **34**: 34.
- 346
- 347 18. Du Z, Xu X, Wu Y, *et al.* COVID-19 serial interval estimates based on confirmed cases  
348 in public reports from 86 Chinese cities. *medRxiv : the preprint server for health sciences*  
349 2020.

350

- 351 19. Zhang J, Litvinova M, Wang W, *et al.* Evolving epidemiology of novel coronavirus  
352 diseases 2019 and possible interruption of local transmission outside Hubei Province in  
353 China: a descriptive and modeling study. *medRxiv : the preprint server for health*  
354 *sciences* 2020.
- 355
- 356 20. Ganyani T, Kremer C, Chen D, *et al.* Estimating the generation interval for coronavirus  
357 disease (COVID-19) based on symptom onset data, March 2020. *Euro surveillance :*  
358 *bulletin Europeen sur les maladies transmissibles = European communicable disease*  
359 *bulletin* 2020; **25**.
- 360
- 361 21. Nishiura H, Linton NM, Akhmetzhanov AR. Serial interval of novel coronavirus  
362 (COVID-19) infections. *International journal of infectious diseases : IJID : official*  
363 *publication of the International Society for Infectious Diseases* 2020; **93**: 284-286.
- 364
- 365 22. Delamater PL, Street EJ, Leslie TF, *et al.* Complexity of the Basic Reproduction Number  
366 (R0). *Emerging infectious diseases* 2019; **25**: 1-4.
- 367
- 368 23. Guo XJ, Zhang H, Zeng YP. Transmissibility of COVID-19 in 11 major cities in China  
369 and its association with temperature and humidity in Beijing, Shanghai, Guangzhou, and  
370 Chengdu. *Infectious diseases of poverty* 2020; **9**: 87.

371

- 372 24. Ran J, Zhao S, Han L, *et al.* A Re-analysis in Exploring the Association between  
373 Temperature and COVID-19 Transmissibility: An Ecological Study with 154 Chinese  
374 Cities. *The European respiratory journal* 2020.  
375
- 376 25. Harmooshi NN, Shirbandi K, Rahim F. Environmental concern regarding the effect of  
377 humidity and temperature on 2019-nCoV survival: fact or fiction. *Environmental science*  
378 *and pollution research international* 2020.  
379
- 380 26. Ren SY, Wang WB, Hao YG, *et al.* Stability and infectivity of coronaviruses in  
381 inanimate environments. *World journal of clinical cases* 2020; **8**: 1391-1399.  
382
- 383 27. Jahangiri M, Jahangiri M, Najafgholipour M. The sensitivity and specificity analyses of  
384 ambient temperature and population size on the transmission rate of the novel  
385 coronavirus (COVID-19) in different provinces of Iran. *The Science of the total*  
386 *environment* 2020; **728**: 138872.  
387
- 388 28. Sahin M. Impact of weather on COVID-19 pandemic in Turkey. *The Science of the total*  
389 *environment* 2020; **728**: 138810.  
390
- 391 29. Zoran MA, Savastru RS, Savastru DM, *et al.* Assessing the relationship between ground  
392 levels of ozone (O<sub>3</sub>) and nitrogen dioxide (NO<sub>2</sub>) with coronavirus (COVID-19) in Milan,  
393 Italy. *The Science of the total environment* 2020; **740**: 140005.  
394

- 395 30. Su D, Chen Y, He K, *et al.* Influence of socio-ecological factors on COVID-19 risk: a  
396 cross-sectional study based on 178 countries/regions worldwide. *medRxiv : the preprint*  
397 *server for health sciences* 2020.
- 398
- 399 31. Bashir MF, Ma B, Bilal, *et al.* Correlation between climate indicators and COVID-19  
400 pandemic in New York, USA. *The Science of the total environment* 2020; **728**: 138835.
- 401
- 402 32. Ahmadi M, Sharifi A, Dorosti S, *et al.* Investigation of effective climatology parameters  
403 on COVID-19 outbreak in Iran. *The Science of the total environment* 2020; **729**: 138705.
- 404
- 405 33. Carraturo F, Del Giudice C, Morelli M, *et al.* Persistence of SARS-CoV-2 in the  
406 environment and COVID-19 transmission risk from environmental matrices and surfaces.  
407 *Environmental pollution* 2020; **265**: 115010.
- 408
- 409 34. Wilson NM, Norton A, Young FP, *et al.* Airborne transmission of severe acute  
410 respiratory syndrome coronavirus-2 to healthcare workers: a narrative review.  
411 *Anaesthesia* 2020; **75**: 1086-1095.
- 412
- 413 35. Faridi S, Niazi S, Sadeghi K, *et al.* A field indoor air measurement of SARS-CoV-2 in  
414 the patient rooms of the largest hospital in Iran. *The Science of the total environment*  
415 2020; **725**: 138401.

416

- 417 36. Sharma AK, Balyan P. Air pollution and COVID-19: Is the connect worth its weight?  
418 *Indian journal of public health* 2020; **64**: S132-S134.  
419
- 420 37. Zhang Y. Dynamic effect analysis of meteorological conditions on air pollution: A case  
421 study from Beijing. *The Science of the total environment* 2019; **684**: 178-185.  
422  
423  
424  
425

426 **Table 1. Multivariable association of meteorological and socio-behavioral covariates with**  
 427 **time dependent R<sub>0</sub> estimates (all study locations, March 1 – May 31, 2020)**

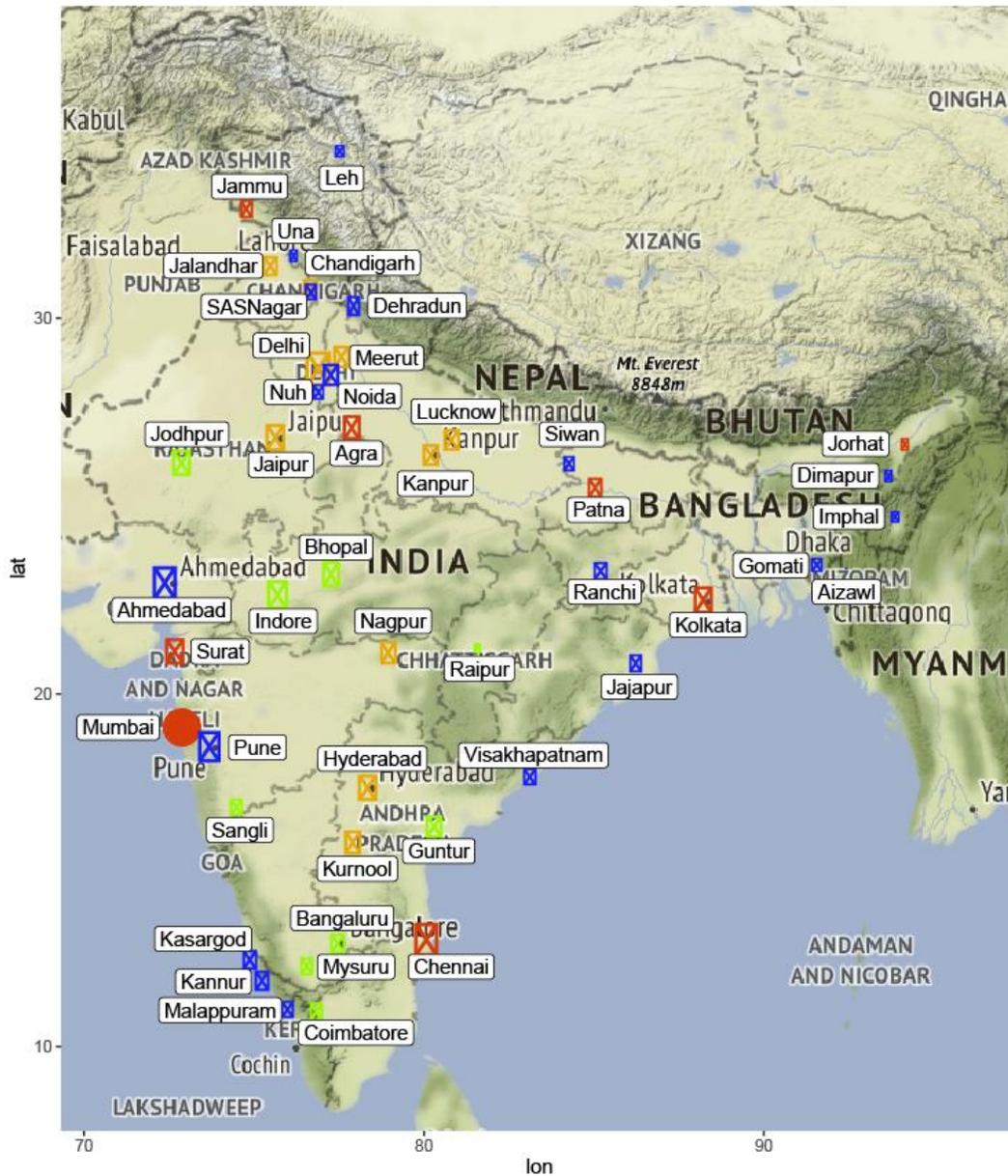
Covariate	$\beta$	95% CI	p
FULL MODEL			
Temperature z-score	-0.10	-0.24 - 0.05	0.190
Relative humidity z-score	-0.04	-0.16 - 0.09	0.539
Air pressure z-score	-0.06	-0.18 - 0.06	0.322
Wind speed z-score	0.06	0.00 - 0.11	0.036
Rainfall z-score	0.02	-0.03 - 0.07	0.471
Retail/recreation z-score	0.54	-0.30 - 1.37	0.204
Grocery/pharmacy z-score	0.17	-0.08 - 0.41	0.184
Parks z-score	-0.26	-0.91 - 0.39	0.426
Transit station z-score	-0.48	-1.18 - 0.23	0.180
Workplaces z-score	0.04	-0.19 - 0.28	0.724
Lockdown phase 1	-0.01	-0.27 - 0.25	0.936
Lockdown phase 2	-0.23	-0.61 - 0.16	0.247
Lockdown phase 3	-0.41	-0.87 - 0.05	0.079
Lockdown phase 4	-0.61	-1.14 - -0.08	0.024
Intercept	1.56	1.30 - 1.81	<0.001
FINAL MODEL			
Temperature z-score	-0.08	-0.13 - -0.03	0.005
Wind speed z-score	0.08	0.03 - 0.12	0.003

Lockdown phase 2	-0.22	-0.34 - -0.09	0.001
Lockdown phase 3	-0.32	-0.45 - -0.19	<0.001
Lockdown phase 4	-0.47	-0.63 - -0.30	<0.001
Intercept	1.51	1.44 - 1.58	<0.001

428  $\beta$ , regression coefficient; CI, confidence interval; p, significance value

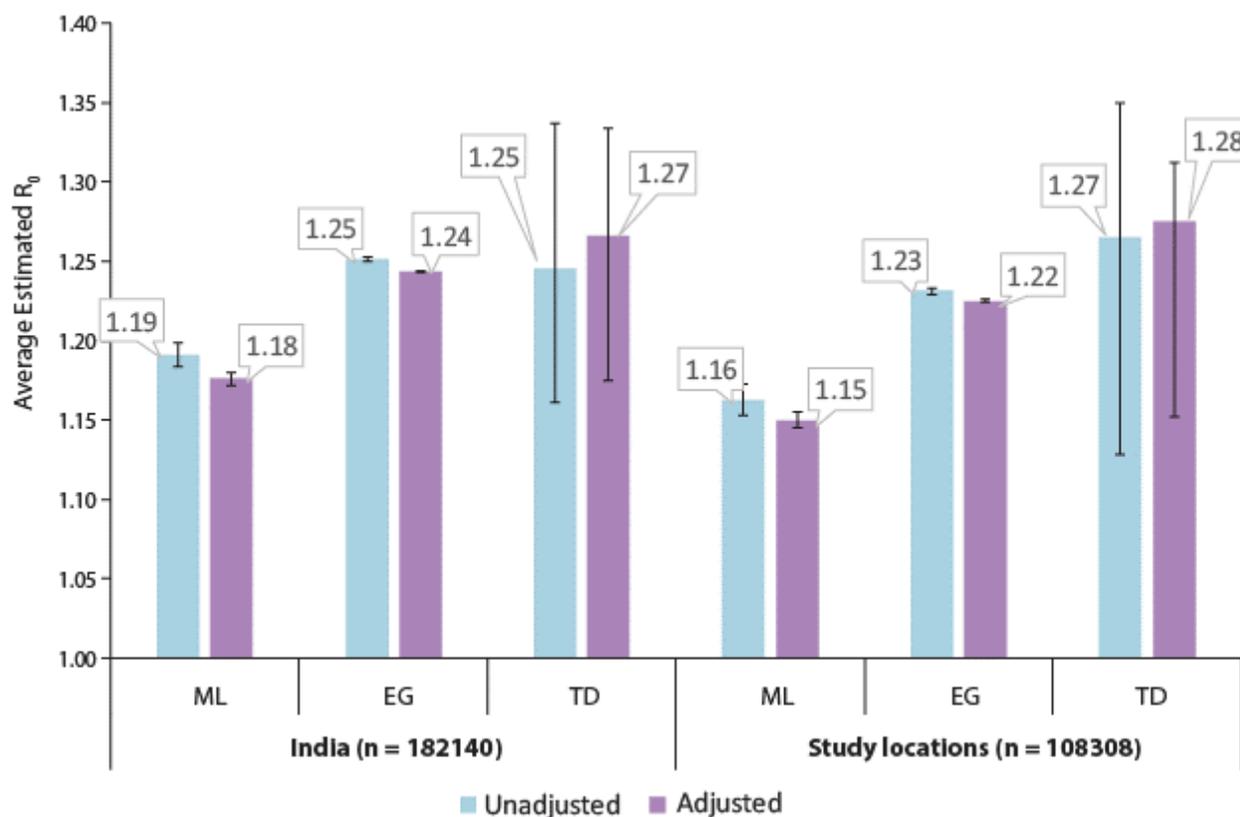
429

430 **Figure 1. Geographical spread, COVID-19 case counts and population density of the study**  
431 **locations.** Selected locations are shown as bubbles, the size of which is proportional to log of  
432 COVID-19 case counts. The color of the bubble indicates quartile of population density based on  
433 the cutoffs mentioned in Supplementary Table 1 – first quartile, blue; second quartile, green,  
434 third quartile orange and fourth quartile, red.



435

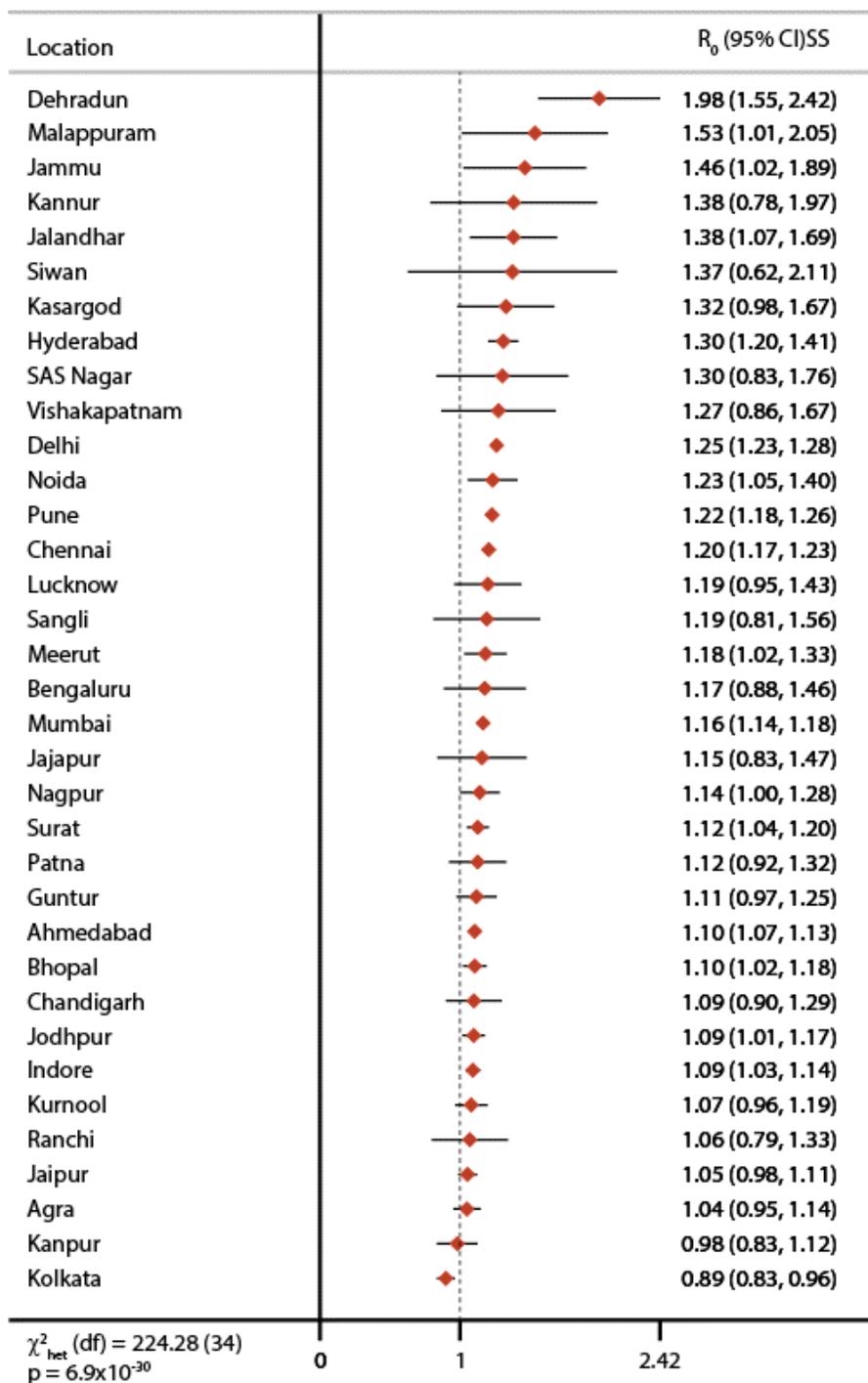
436 **Figure 2. Average estimated  $R_0$  for COVID-19.** Bars show the average  $R_0$  estimates and error  
437 bars indicate the 95% confidence intervals. Average  $R_0$  estimates were derived using three  
438 methods: ML, maximum likelihood; EG, exponential growth; and TD, time dependent. Each  
439 estimate was also derived without adjustment (unadjusted, blue bars) and adjusted for potential  
440 undertesting (adjusted, purple bars).



441

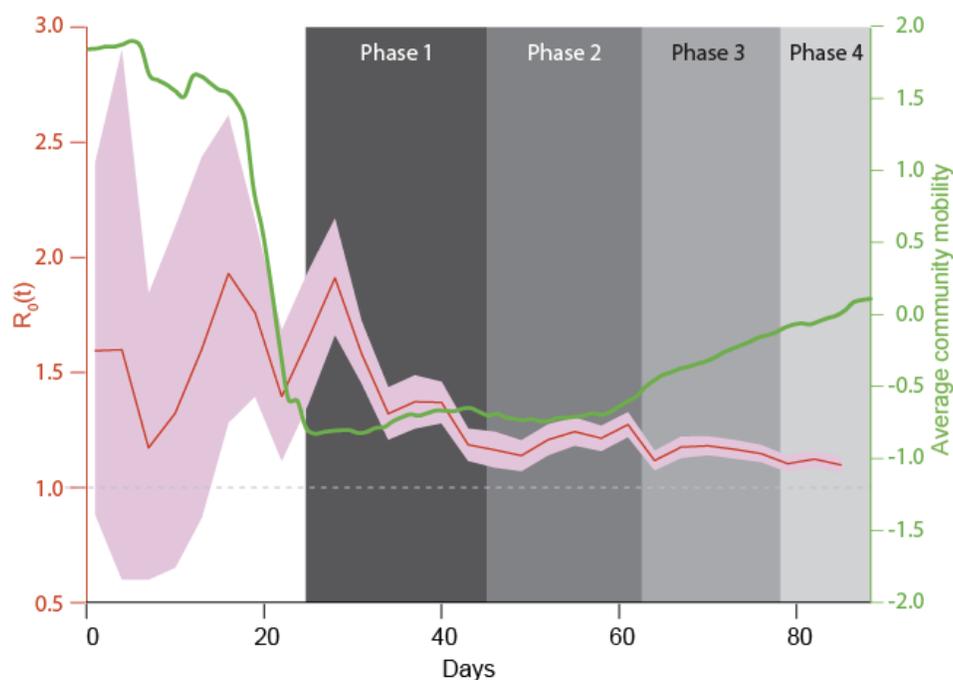
442

443 **Figure 3. Heterogeneity of  $R_0$  estimates across study locations.** The forest plot shows point  
 444 (diamonds) and 95% confidence interval (error bars) estimates for maximum likelihood  
 445 estimates of  $R_0$ .



446

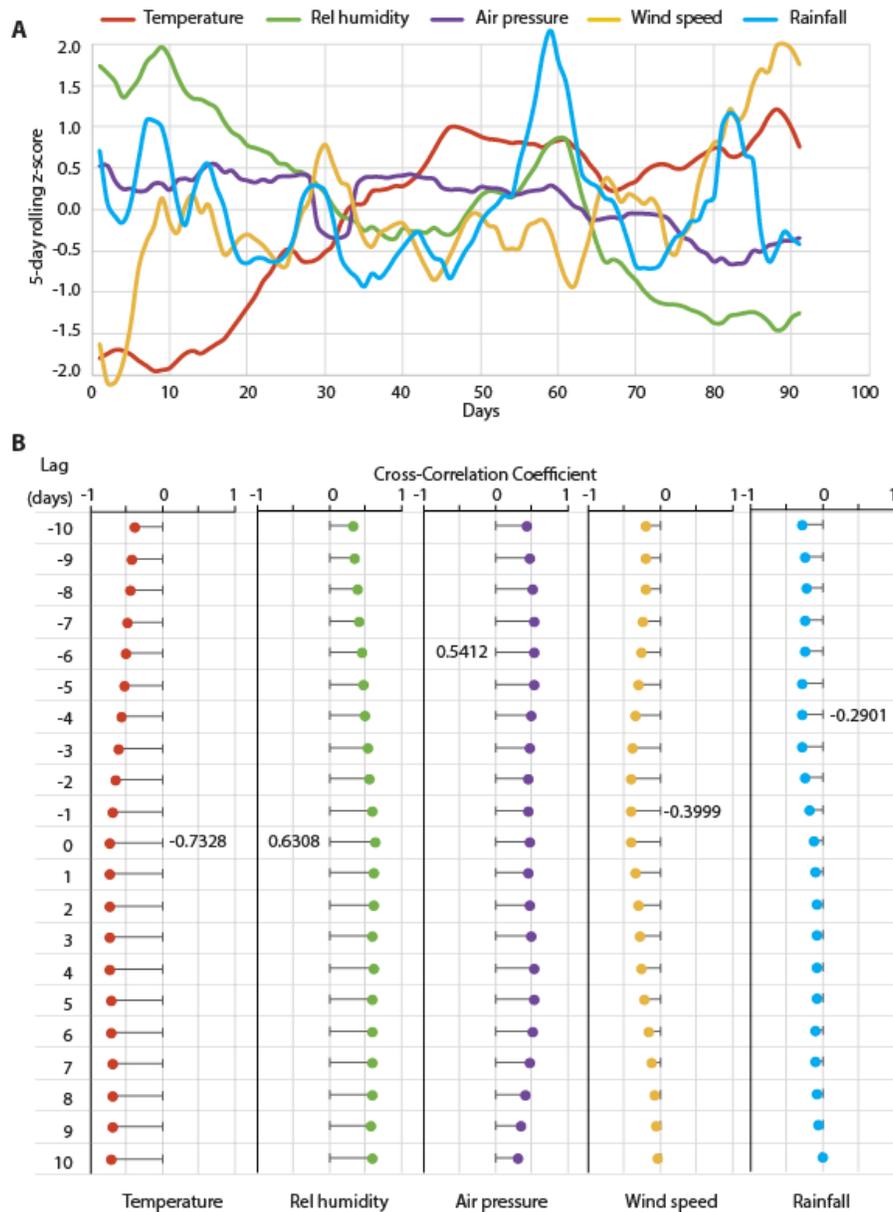
447 **Figure 4. Time dependent  $R_0$  and socio-behavioral interventions.** Red line and pink bands  
448 indicate the time dependent  $R_0$  and 95% confidence intervals, respectively, for each day during  
449 the study. These align to the left axis (colored red). The green curve shows the 5-day rolling  
450 average z-score for cellphone-based mobility data and aligns to the right axis (colored green).  
451 Shaded boxes in the background indicate different phases of the countrywide lockdown in India.



452

453

454 **Figure 5. Meteorological determinants of COVID-19 transmissibility.** (A) Time trends for  
 455 each of the five, color-coded meteorological variables. For each variable, the data were first z-  
 456 transformed and then subjected to a 5-day moving average. (B) Cross-correlograms for  
 457 correlation of each meteorological variable with estimated time dependent  $R_0$ . All cross-  
 458 correlograms were assessed between lags of -10 to 10 days. Most significant correlation for each  
 459 variable is indicated as a number alongside the lag at which it was observed. Rel, relative.



460