Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.
Association of greenness with COVID-19 deaths in India: An ecological study at district level

Ankit Sikarwar a,*, Ritu Rani a,b, Géraldine Duthe a, Valérie Golaz a,c

a French Institute for Demographic Studies (INED), Aubervilliers-Paris, France
b International Institute for Population Sciences, Mumbai, India
c Aix-Marseille University, IRD, LPED, Marseille, France

ARTICLE INFO

Keywords:
Greenness
COVID-19 deaths
NDVI
Environmental health
India

ABSTRACT

Background: The world has witnessed a colossal death toll due to the novel coronavirus disease-2019 (COVID-19). A few environmental epidemiology studies have identified association of environmental factors (air pollution, greenness, temperature, etc.) with COVID-19 incidence and mortality, particularly in developed countries. India, being one of the most severely affected countries by the pandemic, still has a dearth of research exploring the linkages of environment and COVID-19 pandemic.

Objectives: We evaluate whether district-level greenness exposure is associated with a reduced risk of COVID-19 deaths in India.

Methods: We used average normalized difference vegetation index (NDVI) from January to March 2019, derived by Oceansat-2 satellite, to represent district-level greenness exposure. COVID-19 death counts were obtained through May 1, 2021 (around the peak of the second wave) from an open portal: covid19india.org. We used hierarchical generalized negative binomial regressions to check the associations of greenness with COVID-19 death counts. Analyses were adjusted for air pollution (PM$_{2.5}$), temperature, rainfall, population density, proportion of older adults (50 years and above), sex ratio over age 50, proportions of rural population, household overcrowding, materially deprived households, health facilities, and secondary school education.

Results: Our analyses found a significant association between greenness and reduced risk of COVID-19 deaths. Compared to the districts with the lowest NDVI (quintile 1), districts within quintiles 3, 4, and 5 have respectively, around 32% [MRR = 0.68 (95% CI: 0.51, 0.88)], 39% [MRR = 0.61 (95% CI: 0.46, 0.80)], and 47% [MRR = 0.53 (95% CI: 0.40, 0.71)] reduced risk of COVID-19 deaths. The association remains consistent for analyses restricted to districts with a rather good overall death registration (>80%).

Conclusion: Though cause-of-death statistics are limited, we confirm that exposure to greenness was associated with reduced district-level COVID-19 deaths in India. However, material deprivation and air pollution modify this association.

1. Introduction

The worldwide outbreak of COVID-19 has caused a public health emergency and a significant disease burden around the globe. The World Health Organization (WHO) declared the outbreak a pandemic on March 11, 2020, and two years later, there were 437,580,326 cases and 5,977,289 deaths due to coronavirus globally (WHO Coronavirus Dashboard, 2022). As per the official records, a total of 42,931,045 COVID-19 cases and 514,054 related deaths have been reported in India from the onset of the pandemic till March 1, 2022 (WHO Coronavirus Dashboard, 2022). Since the first reported case in February 2020, the Indian government has taken several stringent initiatives (social distancing measures, stay-at-home orders, nationwide lockdowns, etc.) to curb the spread of coronavirus across the nation. Despite those efforts, the pandemic has dramatically hit the country and overwhelmed its health system (Sarkar and Chouhan, 2021). Especially the second wave of COVID-19 struck fiercely, with thousands of daily deaths recorded across the country in April 2021 (Banaji and Gupta, 2021; Ghosh et al., 2021). The pandemic situation in India was severe due to different reasons such as demographic dynamics (population distribution, density, age-sex profile, etc.), poor medical facilities and an intricate socio-economic structure among them (Ghosh et al., 2021; Guilmoto, 2022). The subsequent waves of COVID-19 pandemic have

* Corresponding author. French Institute for Demographic Studies (INED), Campus Condorcet, Aubervilliers-Paris, 93300, France.
E-mail addresses: sikarwar.ankit-kumar@ined.fr, anks.sik@gmail.com (A. Sikarwar).

https://doi.org/10.1016/j.envres.2022.114906
Received 7 June 2022; Received in revised form 18 November 2022; Accepted 21 November 2022
Available online 22 November 2022
0013-9351/© 2022 Elsevier Inc. All rights reserved.
been adversely affecting the health of the population, the economy and the social stability of the country. In this context, it is crucial to identify and understand all the factors that might play a role in preventing the spread of COVID-19 and protecting from related severities.

Environmental factors (air pollution, greenness, wind speed, humidity, temperature, etc.) play a vital role in the occurrence and development of various health outcomes, disease and premature mortality. Several studies have reported the association of COVID-19 with air pollution (Ali and Islam, 2020; Cascetta et al., 2021; Travaglio et al., 2021), temperature, and humidity (Ismail et al., 2022; Khan et al., 2022; Mecenas et al., 2020; Roy, 2021). Exposure to air pollution may also aggravate the COVID-19 severity by negatively affecting individual's immune system (Ali and Islam, 2020; Ismail et al., 2022). In contrast, emerging research on environmental health suggests that the presence of green space is beneficial to the immune system and, more globally, to human health (Engemann et al., 2019; Markeyych et al., 2017; Peng et al., 2020).

At larger spatial scales (such as administrative areas), green space is strongly associated with major environmental parameters such as air pollution, temperature, precipitation, wind, etc. A recent study by Russette et al. (2021) emphasizes the influence of greenspace on reduced risk of COVID-19 deaths through reduced air pollution as well as through improved immune regulation pathways, as greenspace serves as a setting that provides exposure to microbial diversity, such as soil, plants and wildlife. At finer spatial scales (such as neighborhoods), the presence of green spaces may impact COVID-19 infection and mortality as it provides a place for performing physical activity and social interactions while maintaining the recommended safe distance, which is critical for the immune system. Prior studies also suggest socio-economic status as a strong predictor of greenness exposure; low-income neighborhoods being exposed to low levels of greenness and at higher risk of diseases compared to higher socio-economic status neighborhoods (Engemann et al., 2019; Klompmaker et al., 2021). Due to data restrictions, there is no considerable research on neighborhood-level association of greenness and COVID-19. Albeit, selected recent studies in the United States (Kломpmaker et al., 2021; Russette et al., 2021), Canada (Stieb et al., 2020), and Italy (Cascetta et al., 2021; Roviello and Roviello, 2021), have discussed the area-level associations of greenness with COVID-19 incidence. Except for one attempt in urban China (Peng et al., 2022), there is still a dearth of research concerning the role of greenness in reducing COVID-19 death risk, particularly in developing countries. Environmental factors may have a defining and different association with COVID-19 in developing countries characterized by huge population growth, increasing levels of pollution and temperature, urban-rural differences, and a higher proportion of vulnerable population groups, i.e., rural and slum dwellers. Specifically in India, where environmental challenges (Kaur and Pandey, 2021) and socio-economic inequalities (Saikia et al., 2019) have been considerably aggravating human health risks, analyzing the preventing role of greenness in COVID-19 deaths becomes important.

Considering the above-mentioned health benefits of exposure to greenness, we hypothesize that district-level greenness may have positive associations with reduced risk of COVID-19 deaths.

Our goal was to evaluate whether district-level exposure to greenness is associated with a lower risk of COVID-19 deaths in India. To the best of our knowledge, this is the first study to assess the association between greenness exposure and COVID-19 deaths in India. We analyzed data comparing 640 districts organized by NDVI quintiles and adjusted for various district-level indicators. Moreover, we performed various sensitivity analyses to check our results in various circumstances, such as restricted analyses for only districts with better death registration, with higher air pollution levels, with higher shares of materially deprived and affluent households, etc.

2. Data and methods

We used data from multiple sources, including spatial and non-spatial data that are available in the public domain. Table 1 provides detailed information on the outcome variable, explanatory variable, and other covariates used in the analyses. This study considers 640 districts of India as the unit of analysis as classified in the administrative atlas of India in the 2011 Census (Chandramouli, 2011). Data analyses and geoprocessing were performed under STATA 16.0, R studio, and ArcMap 10.3.

2.1. COVID-19 death counts

Data on COVID-19 deaths (total number of reported deaths due to COVID-19 from the onset of the pandemic through May 1, 2021) was collected from the online portal https://www.covid19india.org. This portal compiles COVID-19-related statistics in India using state bulletins and official handles. Though the reporting of COVID-19 deaths is known to be underestimated (Guilmoto, 2022), this source has the strong advantage of covering the whole of India at the district level and has already been used by most of the scientific studies focusing on India (Barman et al., 2020; Dwivedi et al., 2021; Gaur et al., 2021; Kotwal et al., 2020; Kumar and Kumar, 2020; Leffler et al., 2021; Mitra et al., 2020; Sarkar et al., 2020; Tyagi et al., 2021). The timeline of our analysis (from the onset to May 1, 2021) nearly captures COVID-19 deaths up to the pick of the second wave of the pandemic. To explain the association of greenness quintiles with these COVID-19 death counts, we estimated mortality rate ratios (MRRs) by keeping the total population of the district as an offset in the statistical models. MRRs explain the relative difference in COVID-19 deaths associated with successive NDVI quintiles compared to NDVI quintile 1 (group of districts with 0–20% NDVI coverage) (Russette et al., 2021; Stieb et al., 2020).

2.2. Greenness exposure

Normalized difference vegetation index (NDVI) has been widely used in environmental epidemiology, where greenness is linked to various morbidity and mortality patterns (Dadvand et al., 2012; Sadeh et al., 2021; Wang et al., 2019). NDVI values vary from –1 to 1, where values close to one represent areas covered by live green vegetation, values close to zero indicate areas without much vegetation (e.g., rocks, sand) and negative values correspond to various forms of water (Robinson et al., 2017). In our analysis, we used only positive values of NDVI (0 to +1) for estimating the average greenness at district levels and categorized them into equal quintiles. We used continuous raster data of NDVI for India derived by Oceansat-2 satellite and disseminated by the National Remote Sensing Center, India. District-level average NDVI was estimated from the pixel-level (1 km×1 km) NDVI values for January, February, and March 2019. These three months of the year provide NDVI values for the fully grown vegetation without cloud cover for the whole country. Also, during these months, NDVI values are less affected by excesses or deficits of precipitation, which is important for the overall
### Table 1 Description of data used in the study.

| Variable                              | Description                                                                 | Transformation | Source                                                                 |
|---------------------------------------|-----------------------------------------------------------------------------|----------------|------------------------------------------------------------------------|
| COVID-19 deaths (Outcome variable)     | District-level COVID-19 death counts from the onset through May 1, 2021     | None           | Crowdsource data from https://www.covid19india.org/                    |
| Normalized Difference                 | Average greenness index across India (January to March 2019)               | 1 km⁻¹ km spatial grids were rescaled to administrative boundaries of the districts and categorized into quintiles | National Remote Sensing Center, ISRO, Government of India, Hyderabad, India.|
| Vegetation Index (NDVI) (Explanatory variable) | Average annual concentrations of surface PM$_{2.5}$ (2018)               | 0.0$^1$ to 0.1$^1$ spatial grids rescaled to district boundaries | Atmospheric Composition Analysis Group https://sites.wustl.edu/aca/surfacce-pm2-5/ |
| Temperature and Rainfall              | Average annual temperature (2018)                                         | 1 km⁻¹ km spatial grids for all months averaged and rescaled to district boundaries | WorldClim2 https://www.worldclim.org/ |
| Total population                     | Gridded population data (2020)                                            | 100 m¹ 100 m gridded population counts rescaled to and log-transformed | Worldpop https://www.worldpop.org/ |
| Population density                   | Gridded population counts per unit area (2020)                             | 1 km⁻¹ km spatial grids of population density averaged and rescaled for district boundaries | Census of India 2011 https://cemensusindia.gov.in/2011-common/censusdata2011.html |
| Proportion of older adults (50 years and above) and their sex ratio (males per 100 females) | Gridded population counts of males and females above 50 years (2020)         | Age and sex groups of 100 m² 100 m gridded population (50+) counts added for district boundaries | Census of India 2011 https://cemensusindia.gov.in/2011-common/censusdata2011.html |
| Rural population                     | Proportion of rural population (2011)                                      | None           | Census of India 2011 https://cemensusindia.gov.in/2011-common/censusdata2011.html |
| Household crowding                   | Percentage of HH having 3 or more persons per sleeping room (2016)         | None           | National Family Health Survey – 4 (NFHS-4) http://rchiips.org/nfhs4/nfhs4.html |
| Material deprivation                 | Households having at most one of the following eight assets or amenities - pucca house, electricity connection, phone, television, AC/cooler, refrigerator, washing machine, and motorized vehicle-have been defined as materially deprived households. Whereafter | None           | National Family Health Survey – 4 (NFHS-4) http://rchiips.org/nfhs4/nfhs4.html |

### Table 1 continued

| Variable                              | Description                                                                 | Transformation | Source                                                                 |
|---------------------------------------|-----------------------------------------------------------------------------|----------------|------------------------------------------------------------------------|
| Household crowding                   | Percentage of HH having 3 or more persons per sleeping room (2016)          | None           | National Family Health Survey – 4 (NFHS-4) http://rchiips.org/nfhs4/nfhs4.html |
| Material deprivation                 | Households having at most one of the following eight assets or amenities - pucca house, electricity connection, phone, television, AC/cooler, refrigerator, washing machine, and motorized vehicle-have been defined as materially deprived households. Whereafter | None           | National Family Health Survey – 4 (NFHS-4) http://rchiips.org/nfhs4/nfhs4.html |

We estimated district-level average annual concentrations of surface PM$_{2.5}$ for 2018 using the data obtained from the Atmospheric Composition Analysis Group (Hammer et al., 2020) with a resolution of 1 km. Other environmental factors: temperature and rainfall were estimated from the worldclim2 data portal (Fick and Hijmans, 2017) by averaging the monthly spatial layers for 2018. The latest population counts, proportion of population aged 50 years and more (% older adults), number of males per 100 females of population aged 50 years and more (sex ratio over age 50), and population density was aggregated from the gridded population data of 2020 produced by Worldpop (WorldPop, 2018). Fig. 1 illustrates the raster surfaces of spatial parameters used in the study. District estimates of rural population (% rural) were derived from the 2011 Census of India (Chandramouli and General, 2011). District-level household characteristics such as household crowding (HH crowding), proportion of materially deprived households (% materially deprived HH), and proportion of households having a member with at least 10 years of education (secondary school education) were estimated from the National Family Health Survey – 4 conducted in 2016 (IIPS, 2017). The number of government healthcare infrastructures (health facilities) was derived from the Rural Health Statistic Report 2020-21 (Government of India, 2021).

#### 2.3. Covariates

Covariates included variables that could potentially confound the association of greenness and the risk of COVID-19 deaths based on the evidence from previous research (Bowe et al., 2021; Cascetta et al., 2021; Lee et al., 2021; Lorenzo et al., 2021; Lu et al., 2021; Russette et al., 2021; Stieb et al., 2020; Travaglio et al., 2021). We utilized the most recent estimates for the better reliability of our results (Table 1). We estimated district-level average annual concentrations of surface PM$_{2.5}$ for 2018 using the data obtained from the Atmospheric Composition Analysis Group (Hammer et al., 2020) with a resolution of 1 km. Other environmental factors: temperature and rainfall were estimated from the worldclim2 data portal (Fick and Hijmans, 2017) by averaging the monthly spatial layers for 2018. The latest population counts, proportion of population aged 50 years and more (% older adults), number of males per 100 females of population aged 50 years and more (sex ratio over age 50), and population density was aggregated from the gridded population data of 2020 produced by Worldpop (WorldPop, 2018). Fig. 1 illustrates the raster surfaces of spatial parameters used in the study. District estimates of rural population (% rural) were derived from the 2011 Census of India (Chandramouli and General, 2011). District-level household characteristics such as household crowding (HH crowding), proportion of materially deprived households (% materially deprived HH), and proportion of households having a member with at least 10 years of education (secondary school education) were estimated from the National Family Health Survey – 4 conducted in 2016 (IIPS, 2017). The number of government healthcare infrastructures (health facilities) was derived from the Rural Health Statistic Report 2020-21 (Government of India, 2021).

#### 2.4. Statistical analysis

We used a generalized negative binomial regression model to produce MRRs, enabling us to evaluate the associations between greenness and COVID-19 deaths. To examine the effects of potential confounding factors, we constructed a series of models with increasing covariate adjustment. Model 1 only included NDVI quintiles and a population size offset. In model 2, we additionally adjusted for important environmental covariates: PM$_{2.5}$, temperature, and rainfall. In model 3, we added all district-level demographic and household covariates: population...
density, % older adults, sex ratio over age 50, % rural, HH crowding, % materially deprived HH, health facilities, and secondary school education.

To check possible collinearity among predictors within our full model, we estimated variance inflation factors (VIF) for all covariates. In general, VIF values of 5 or greater present a problem of collinearity (Hair et al., 2011). In our analysis, all VIFs were below 2.5 (Table A1), indicating covariates without serious collinearity. Before the main analysis, a bivariate local Moran’s index was calculated to check patterns of spatial autocorrelation between COVID-19 deaths and NDVI.

We performed sensitivity analyses to evaluate the magnitude of change in the relationship between greenness and COVID-19 deaths by considering various influential criteria described below. We considered NDVI as a continuous variable in the sensitivity analyses for simplicity.

- According to experts, COVID-19 death statistics in India remained dramatically under-reported due to various reasons (Guilmoto, 2022; Ioannidis, 2021; The Lancet, 2021). Keeping in mind the huge under-reporting of deaths in India, we identified the districts of India with more than 80 percent of death registration during the three years preceding the survey according to NFHS-5, a national survey conducted in 2019–21 (IIPS, 2021). We believe that the COVID-19 deaths reported in these specific districts provide a picture relatively close to the actual COVID-19 deaths that occurred. We performed a separate analysis to check the reliability of our results only in those districts.

- Area-level exposures to higher levels of PM$_{2.5}$ concentrations may aggravate the risk of COVID-19 (Klopfmacher et al., 2021) and may change its association with greenness. Thus, we conducted a separate analysis of districts with PM$_{2.5}$ concentrations higher than 40 μg/m$^3$.

- Socio-economic inequalities may present a differential mortality risk for population (Saikia et al., 2019). Thus, we performed separate analyses for districts with a higher share of materially deprived households (>20%) and affluent households (>40%).

- Pronounced differences between urban and rural regions in COVID-19 deaths and infections, coverage and accuracy of reporting, availability, access to healthcare system, environmental conditions, etc., could indirectly impact the linkages of greenness and COVID-19 deaths. We also performed separate analyses for ‘rural’ districts (i.e., districts where over 80% of the population is rural) and ‘urban’ districts (districts where over 60% of the population is urban).

- Lastly, we also checked the association using a multilevel model considering State as random effect.

3. Results

Fig. 2 displays district-wise spatial patterns of NDVI and COVID-19 deaths (per 1000 people) in India. NDVI and COVID-19 deaths were spatially varying across Indian districts. The bivariate spatial autocorrelation (Moran’s I = $-0.029$, 95% CI) also suggests spatial heterogeneity in the association of NDVI and COVID-19 deaths (Fig. A1). NDVI values were higher for the districts in North-central, Western coastal, and North-eastern parts of India. COVID-19 death statistics were more severe in the districts including big urban centers of India such as Mumbai, Delhi, Kolkata, Chennai, Bangalore, etc. Almost 20% of the total COVID-19 deaths till May 1, 2021, were concentrated in top 2% of the urban districts. Table 2 presents summary statistics at district level for the variables used in the models. Mean COVID-19 case fatality of 1.0% suggests that, on an average, for every 100 cases of COVID-19 there was one death.

1 The chosen thresholds provide a sufficient number of districts in each category in order to conduct the analysis.
We employed three generalized negative binomial regression models to check associations of COVID-19 deaths and NDVI quintiles with different covariates (Table 3). Mortality rate ratios (MRRs) for NDVI quintiles from three models are illustrated in Fig. 3. MRR values in model 1 make no adjustments for the effects of the model predictors. The second model presents MRR values adjusted for three additional environmental covariates (PM$_{2.5}$, temperature, and rainfall). In the main model (model 3), MRR values are fully adjusted for all the predictors used in the analysis. Compared to the 1st quintile of districts with lowest NDVI, districts characterized with a higher NDVI in quintiles 3, 4, and 5 have respectively, around 32% [MRR = 0.68 (95% CI: 0.51, 0.88)], 39% [MRR = 0.61 (95% CI: 0.46, 0.80)], and 47% [MRR = 0.53 (95% CI: 0.40, 0.71)] reduced risk of COVID-19 deaths.

As per the fully adjusted model 3 (main model), some of the covariates showed a significant association with COVID-19 deaths. Districts with higher sex ratio of older adults [MRR = 0.98 (95% CI: 0.96, 0.98)] and share of rural population [MRR = 0.97 (95% CI: 0.97, 0.98)] have relatively significantly reduced risk of COVID-19 deaths (2% and 3% respectively). Whereas districts with increasing PM$_{2.5}$ [MRR = 1.005...

---

**Table 2**

Summary statistics of key indicators at district level (Sources: refer Table 1).

| Variables                     | Mean   | Std. Dev. | Minimum | Maximum |
|-------------------------------|--------|-----------|---------|---------|
| Population (n)                | 1,997,456 | 1,668,682 | 9243    | 15,300,000 |
| COVID-19 deaths (n)           | 339    | 836       | 0       | 9642    |
| Case Fatality Ratio (%)       | 1.0    | 0.6       | 0       | 3.9     |
| NDVI                          | 0.5    | 0.1       | 0.1     | 0.7     |
| PM$_{2.5}$ (μg/m$^3$)         | 68.2   | 31.2      | 0.0     | 149.4   |
| Temperature (°C)              | 27.5   | 5.9       | -5.6    | 33.2    |
| Rainfall (cm)                 | 100.0  | 54.4      | 8.8     | 408.9   |
| Population density (person per km$^2$) | 964 | 3227      | 1       | 48,246 |
| % older adults (%)            | 19.3   | 4.1       | 4.1     | 48.2    |
| Sex ratio over age 50 (male per 100 female) | 99.0 | 9.2       | 74      | 143     |
| % rural (%)                   | 73.6   | 21.2      | 0.0     | 100.0   |
| HH crowding (%)               | 39.7   | 15.9      | 2.4     | 75.1    |
| % materially deprived HH (%)  | 12.2   | 13.0      | 0       | 61.1    |
| Health facilities (n)         | 282    | 193       | 5       | 1184    |
| Secondary school education (%)| 46     | 11        | 18      | 100     |

**Table 3**

Association between greenness, selected covariates and COVID-19 deaths using negative binomial regression (N = 640) (Sources: refer Table 1).

| Variables             | Model 1 | Model 2 | Model 3 |
|-----------------------|---------|---------|---------|
|                       | MRR (CI)| MRR (CI)| MRR (CI)|
| COVID-19 deaths       |         |         |         |
| Greenness (NDVI)      |         |         |         |
| Quintile 1            |         |         |         |
| Quintile 2            | 0.653* (0.47, 0.90) | 0.557*** (0.40, 0.77) | 0.844 (0.65, 1.09) |
| Quintile 3            | 0.409*** (0.30, 0.57) | 0.329*** (0.24, 0.46) | 0.676** (0.51, 0.88) |
| Quintile 4            | 0.402*** (0.29, 0.56) | 0.303*** (0.21, 0.43) | 0.610*** (0.46, 0.80) |
| Quintile 5            | 0.464*** (0.33, 0.64) | 0.370*** (0.26, 0.52) | 0.529*** (0.40, 0.71) |
| PM$_{2.5}$            | 1.005** (1.00, 1.01) | 1.005* (1.00, 1.01) | 1.005 (1.00, 1.01) |
| Temperature           | 1.060*** (1.04, 1.08) | 0.998 (0.98, 1.02) | 1.000 (0.99, 1.00) |
| Rainfall              | 1.003* (1.00, 1.01) | 0.999 (0.99, 1.00) | 1.000 (0.99, 1.00) |
| Population density    | 1.000 (0.99, 1.00) | 1.000 (0.99, 1.00) | 1.000 (0.99, 1.00) |
| % older adults (%)     | 1.029* (1.01, 1.05) | 0.976*** (0.96, 0.99) | 0.976*** (0.97, 0.98) |
| Sex ratio over age 50 | 1.000 (0.99, 1.01) | 1.000 (0.99, 1.01) | 1.000 (0.99, 1.01) |
| % rural (%)           | 0.974*** (0.97, 0.98) | 1.000 (0.99, 1.01) | 1.000 (0.99, 1.01) |
| HH crowding (%)       | 1.000 (0.99, 1.01) | 1.000 (0.99, 1.01) | 1.000 (0.99, 1.01) |
| % materially deprived HH | 1.000 (0.99, 1.01) | 1.000 (0.99, 1.01) | 1.000 (0.99, 1.01) |
| Health facilities (%) | 1.000 (0.99, 1.01) | 1.000 (0.99, 1.01) | 1.000 (0.99, 1.01) |
| Secondary school education | 1.000 (0.99, 1.01) | 1.000 (0.99, 1.01) | 1.000 (0.99, 1.01) |
| Constant              | 0.952 (0.76, 1.19) | 0.112*** (0.07, 0.19) | 0.883*** (0.03, 0.22) |
| Log (population)      | 1 (offset) | 1 (offset) | 1 (offset) |

95% confidence intervals in brackets. *p < 0.05, **p < 0.01, ***p < 0.001. Reference category, MRR: Mortality rate ratio, CI: confidence interval, HH: households.
(95% CI: 1.00, 1.01)), proportion of older adults [MRR = 1.029 (95% CI: 1.01, 1.05)], health facilities [MRR = 1.002 (95% CI: 1.00, 1.00)], and secondary school education [MRR = 1.045 (95% CI: 1.03, 1.06)] were found to have significant and slightly higher risk of COVID-19 deaths (0.05%, 3%, 0.02%, and 4% respectively). No significant associations were found for COVID-19 death with temperature, rainfall, population density, household crowding, and materially deprived households. Although in the partially adjusted model 2, along with PM$_{2.5}$, temperature and rainfall showed a significant and positive association (6% and 0.3% increase in MRR, respectively) with COVID-19 deaths.

Results of sensitivity analyses are presented in Table 4. Compared with the results of the main model (MRR = 0.164), associations of average NDVI with the risk of COVID-19 deaths remained almost unchanged (MRR = 0.178) in case of the districts having a relatively higher death registration coverage (>80%) during previous three years of the NFHS-5 (2019–21) survey. Whereas, for the districts with higher levels of PM$_{2.5}$ (>40 μg/m$^3$) and higher shares of materially deprived households (>20%) the positive associations between greenness and reduced risk of COVID-19 deaths (MRR = 0.31 and 0.69, respectively) were found weaker compared to the main model. However, our results remained insignificant for the districts with high shares of affluent households and urban population, and also when applying multilevel analysis considering State as random effect.

### 4. Discussion

In this study, we found that greenness was inversely associated with COVID-19 deaths in the districts of India. The results indicated a lower risk of COVID-19 deaths in the districts with higher levels of average NDVI values. The associations remained consistent in models restricted for critical criteria in sensitivity analyses. Extending on the ecological approach used by Russette et al. (2021) and Klompmaker et al. (2021) in their county-level studies in the United States, our study finds similar results showing the potential protective effects of greenness exposure on COVID-19 deaths in India. Similar positive associations between greenness and reduced risk of COVID-19 were evident in a few recent ecological studies (Cascetta et al., 2021; Peng et al., 2022; Roviello and Roviello, 2021; Stieb et al., 2020). Except for a recent study in China (Peng et al., 2022), all previous studies linking greenness with COVID-19 risk were focused on developed countries where reliable data on COVID-19 and environmental factors are available in the public domain (Kломpmaker et al., 2021; Lu et al., 2021; Russette et al., 2021; Sadieh, 2021; Spotswood et al., 2021). Extending on the arguments and hypothesis of these research works, our study used up-to-date population data derived from gridded spatial data and also tested for a possible incompleteness of death reporting in India.

While ecological approaches, including ours, are not suitable to establish strong causal relationships, previous research suggests multiple possible mechanisms that could explain the statistical association of greenness with COVID-19 risk. Living in greener areas may impact human health in various ways such as better mental health (Crouse et al., 2021; Weimann et al., 2015), improved immunity (Fuertes et al., 2014; Gascon et al., 2016), reduced respiratory and cardiovascular risk (Eldeirawi et al., 2019; Yeager et al., 2020), higher physical activity (Myttton et al., 2012; Richardson et al., 2013; Ward et al., 2016). Placing our results in context, previous studies conducting exposure assessment in environmental epidemiology have found a significant association of unit increase in surrounding NDVI with various health outcomes. For instance, greenness has been found to be associated with psychiatric disorders (Engemann et al., 2019), child birth weight (Cusack et al., 2017), female mortality (James et al., 2016), cardiovascular risk factors (Brown et al., 2016), and physical activity (McMorris et al., 2015). A study by (Spotswood et al., 2021) highlights positive impacts of greenness exposure to keep the COVID-19 infected cases subclinical or asymptomatic by boosting the natural killer (NK) cells in the human body (Adiba and Jaelani, 2021; Xu et al., 2021) which are proven to

---

### Table 4

Summary of sensitivity analyses of the association of average NDVI and COVID-19 deaths (Sources: refer Table 1).

| Model                                      | MRR   | p-value | 95% CI    | N  |
|--------------------------------------------|-------|---------|-----------|----|
| Main model$^a$                            | 0.164 | 0.00    | (0.073, 0.365) | 640 |
| For districts with improved death registration$^b$ | 0.178 | 0.01    | (0.045, 0.709) | 249 |
| For districts with high PM$_{2.5}$ levels$^c$ | 0.309 | 0.01    | (0.126, 0.756) | 530 |
| For districts with high proportion of materially deprived HH$^d$ | 0.688 | 0.01    | (0.009, 0.531) | 152 |
| For districts with high proportion of affluent HH$^e$  | 1.314 | 0.78    | (0.164, 10.48) | 124 |
| For ‘rural’ districts$^f$                  | 0.096 | 0.00    | (0.035, 0.260) | 318 |
| For ‘urban’ districts$^g$                 | 0.488 | 0.72    | (0.010, 23.35) | 53  |
| Multilevel model (State as random effect) | 0.908 | 0.83    | (0.375, 2.196) | 640 |

$^a$ Similar with model 3 in the main analysis, except NDVI is continuous here.
$^b$ Death registration above 80%.
$^c$ Average PM$_{2.5}$ higher than 40 μg/m$^3$.
$^d$ Proportion of materially deprived households >20%.
$^e$ Proportion of affluent households >40%.
$^f$ Proportion of rural population more than 80%.
$^g$ Proportion of urban population more than 60%.
fight with the virus-infected cells (Björkström et al., 2022; Market et al., 2020). Greener areas also play a crucial role in controlling the levels of air pollution (Lei et al., 2021) and temperature (Vaz Monteiro et al., 2016), which have emerged as two of the most influential environmental predictors of higher COVID-19 incidences and mortality across the world (Bowe et al., 2021; Cascetta et al., 2021; Lorenzo et al., 2021; Mecenas et al., 2020; Roy, 2021; Stieb et al., 2020). Our results (model 2, Table 3) also suggest a higher risk of COVID-19 deaths in the districts with higher PM2.5 levels and temperature.

In our analysis, the proportion of population aged 50 years and above was positively associated with COVID-19 deaths. At the individual level, a combination of reduced immune systems, limited physical activities, and existing comorbidities among older adults place them at a greater risk for health complications and death as a consequence of COVID-19, as highlighted by Hashim et al. (2020) and Yanez et al. (2020) in their international comparisons. Furthermore, we found a negative association between COVID-19 deaths and the sex ratio of older adults: districts with a higher proportion of older women experience higher risks of COVID-19 deaths. Deviated from the global trend, sex differences in COVID-19 death risk might be due to discrepant COVID-19 data and biases in COVID-19 case identification by sex (Odehigia and Raj, 2021). It is also evident that, in India, there exists a clear gender gap in hospitalization rates (Kumar et al., 2020) as well as significant gender bias in the utilization and financing of elderly inpatient care (Joe et al., 2017). We also found that districts with a higher share of the rural population were at lower risk of COVID-19 deaths. It might be mainly due to the higher number of unreported COVID-19 deaths in rural areas of India (Zimmermann et al., 2021). It may also be possible due to the higher greenness levels, the lower levels of air pollution, and population density in rural areas. In addition to this, risk factors such as hypertension, diabetes, obesity, etc., have been more prevalent in urban areas compared to rural areas of India (Geldsetzer et al., 2018; Pandey et al., 2021; Sinha et al., 2022) and this makes urban areas more exposed to COVID-19 deaths. The association of population density with COVID-19 deaths remained insignificant in our analysis. This inconsistent associations of population density with COVID-19 cases and deaths were also evident in other studies using similar approaches (Kloppingaker et al., 2021; Peng et al., 2020; Spotswood et al., 2021). Considering the fact that these studies have been done at bigger spatial scales, population density might be an important covariate in assessing the linkages of environment and COVID-19 risk at finer spatial scales. Our sensitivity analyses provided some important insights into the changing association of greenness with COVID-19 deaths according to different district-level conditions. There was a marginal deviation in MRR values, when we restricted our analysis to the districts with higher levels of death registration. This implies consistency for our model even in case of incomplete death registration. Moreover, as discussed earlier, the existing association of higher NDVI with lower risks of COVID-19 deaths was significantly reduced in the districts with higher levels of air pollution (PM2.5). More importantly, the presence of poor households in a district significantly reduces this positive association between greenness and lower risks of COVID-19 deaths. The extremely low levels of MRR for 100% rural districts again hint towards the huge underreporting of COVID-19 deaths. The insignificant results of the multilevel model, considering State as random effect, are supported by the spatial autocorrelation statistics indicating spatial heterogeneity in the associations of greenness and COVID-19 deaths. Also, in India, at least until the second wave of the pandemic, major interventions were designed at the national level following a classification of the districts by severity (red, green, and orange zones).

Our study employed ecological regression models that present a simple and cost-effective approach capable of assessing potential associations between greenness exposure and COVID-19 deaths in a vast spatial setting with a huge population. Besides this methodological aspect, there are three major strengths in our study. First, to our understanding, this is the earliest attempt to use the most recent population estimates derived from gridded data as an input to link greenness with COVID-19 deaths. Second, our analysis captures the situation during the second wave of the pandemic, which accounted for maximum deaths in India. Third, keeping in mind the under-reporting of COVID-19 deaths in India, we did a separate analysis for the districts with improved death reporting percentages. Our study can be insightful for developing countries where the association of COVID-19 with environmental factors could be strong, though it remained unexplored.

We acknowledge that this study has several limitations. As adopted by previous studies (Cascetta et al., 2021; Peng et al., 2022; Russette et al., 2021; Spotswood et al., 2021), we used an ecological approach to generate our findings at district levels due to the fact that COVID-19 data are not available at individual-level. Area-level analyses are unable to adjust for multiple individual-level risk factors (e.g., existing morbidities, age, socio-economic conditions, smoking status, etc.), and the results might be affected by ecological fallacy (Peng et al., 2022; Spotswood et al., 2021). Such results could be used for hypothesis-generating motives rather than making causal inferences (Kloppingaker et al., 2021). Another limitation of our study is the reliability of COVID-19 mortality data that is claimed to be significantly under-reported in India (Guilmoto, 2022; Li and Nair, 2021). There is a strong need for accurate data on COVID-19 at smaller spatial scales and at individual levels so that environmental factors can be addressed at these finer scales. Although the analyses were based on the most recent data on various parameters, a difference exists in their acquisition time. It would be important to note that the aggregated environmental parameters at the district level may not change much over a short period. Undoubtedly, the availability of up-to-date socio-demographic parameters have influenced the model results. Moreover, remotely sensed data makes estimation of greenness possible with different vegetation indices, but doing so, the details about the composition of green spaces (such as playgrounds, parks, forests, farms, etc.) are often unidentifiable, particularly when looking at larger spatial scales such as districts, states, and countries. Since the analyses were conducted at district level, the linkages of greenspace with COVID-19 deaths in rural and urban places could not be captured, though we included the proportion of rural and urban population in our analyses. It would be interesting, if data were available at that level, to check these associations for villages on one hand, and for cities, which are still smaller than districts, on the other, as green space interacts with health in different ways in rural and urban settings.

5. Conclusion

We performed analyses to link greenness with COVID-19 deaths in Indian districts until the peak of the second wave of the pandemic. After controlling for selected covariates, exposure to greenness exhibited a positive association with a reduced risk of COVID-19 deaths, confirming the findings of the earliest environmental epidemiology studies on developed countries. The positive association of greenness with COVID-19 deaths can be modified according to various environmental, demographic and socio-economic conditions. Our study also emphasized the role of gridded population estimates which can be alternatively used if the census data are not available during the time of the pandemic. Though our results are not suitable for producing causal relationships, they can be used for hypothesis building for future research when individual-level or finer-scale data on COVID-19 will be publicly available. At the same time, our findings imply that living in greener areas may have significant public health benefits, which are crucial to inform policy for controlling and preventing the adversities associated with COVID-19 and with future pandemics.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.
Credit author statement

Ankit Sikarwar: Conceptualization, Methodology, Validation, Formal Analysis, Writing – original draft, Writing – review & editing, Visualization, Ritu Rani: Conceptualization, Methodology, Formal Analysis, Writing – original draft, Writing – review & editing, Geraldine Duthe: Methodology, Validation, Formal Analysis, Writing – review & editing, Valerie Golaz: Methodology, Validation, Formal Analysis, Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Appendix

Fig. A. 1. Bivariate spatial autocorrelation between COVID-19 deaths (per 1000 people) and NDVI (Source: refer Table 1).

Table A.1

| Variables                        | VIF  |
|----------------------------------|------|
| NDVI quintiles                   |      |
| 2                                | 1.87 |
| 3                                | 1.96 |
| 4                                | 2.2  |
| 5                                | 2.32 |
| PM2.5                            | 2.12 |
| Temperature                      | 2.2  |
| Rainfall                         | 1.78 |
| Population density               | 1.62 |
| Population aged 50 years and above (%) | 1.34 |
| 50+ sex ratio                    | 1.6  |
| Rural population                 | 2.14 |
| HH crowding                      | 1.97 |
| Materially deprived HH           | 2.41 |
| Health facility                  | 1.3  |
| Secondary school education       | 2.23 |
| Mean VIF                         | 1.98 |
