

**The COVID-19 pandemic's footprint in India:
An assessment of the district-level susceptibility and
vulnerability**

**Kashif Imdad, Meheeb Sahana, Md Juel Rana, Ismail Haque, Priyank Pravin Patel,
Malay Pramanik, Preeti Dhillon, Laxmi Kant Dwivedi**

A Situational Analysis Paper for Policy Makers



International Institute for Population Sciences, Mumbai

(www.iipsindia.ac.in)

June 1, 2020

Paper 17: The COVID-19 pandemic's footprint in India: An assessment on the district-level susceptibility and vulnerability

(This work has not been peer-reviewed. If you find any errors or have clarifications, please email the authors)

For research and updates on Covid-19, visit:

<https://www.iipsindia.ac.in/content/covid-19-information>

Team of Researchers

Md Juel Rana, Preeti Dhillon, and Laxmi Kant Dwivedi

International Institute for Population Sciences, Mumbai - 400088, India

Email: jranajnu@gmail.com; pdhillon.iips@gmail.com; laxmikdwivedi@gmail.com

Kashif Imdad

Department of Geography, Pandit Prithi Nath PG College (affiliated to Chhatrapati Shahu Ji Maharaj University), 96/12, Mahatma Gandhi Marg, Kanpur - 208001, Uttar Pradesh, India

Email: kashifimdad@gmail.com

Meheebub Sahana

School of Environment, Education and Development, University of Manchester, Oxford Road, Manchester M13 9PL, United Kingdom

Email: mehebubsahana@gmail.com

Ismail Haque

Indian Council for Research on International Economic Relations (ICRIER)

Plot No. 16-17, Sector-6, Pushp Vihar Institutional Area, Saket, New Delhi - 110017, India

and

Centre for the Study of Regional Development, School of Social Sciences, Jawaharlal Nehru University, New Delhi - 110067, India

Email: ismailhaque140489@gmail.com

Priyank Pravin Patel

Department of Geography, Presidency University, 86/1, College Street, Kolkata - 700073, West Bengal, India

Email: priyank999@hotmail.com

Malay Pramanik

Department of Development and Sustainability, School of Environment, Resources and Development, Asian Institute of Technology (AIT), PO. Box 4, Klong Luang, Pathumthani - 12120, Thailand.

and

Centre of International Politics, Organization, and Disarmament, School of International Studies, Jawaharlal Nehru University, New Delhi - 110067, India

Email: malaygeo@gmail.com

The COVID-19 pandemic's footprint in India: An assessment of the district-level susceptibility and vulnerability

Abstract

In this nationwide study, we trace the COVID-19 pandemic's footprint across India's districts. We identify its primary epicentres as the international airports of Mumbai and Delhi and track the outbreak into India's hinterlands in separate time-steps that encapsulate the different lockdown stages implemented. Using a detailed district-level database encompassing climatic, demographic and socioeconomic parameters, we identify hotspots and significant clusters of COVID-19 cases, discerning temporal changes and highlighting areas where the pandemic may spread next. Of prime concern are significant clusters in the country's west and north parts and rising numbers in the east as lockdown measures ease, migrant workers return home and the economy resumes. Exemplar insights emerge from Kerala in South India, where virus hotspots were minimised though effective contact-tracing, mass testing and accessible treatment though rebounding back sharply as the nation restarts. We also perform an epidemiological and socioeconomic susceptibility and vulnerability analyses. The former elicits areas whose populations may possibly be physiologically weaker in combating the virus and therein we expect a high incidence of cases. The latter shows regions that can report high fatalities due to ambient poor demographic and health-related factors. Correlations derived from the generalised additive model show that a high share of urban population and high population density (1500-2500 people/km²), particularly in slum areas, elevate the COVID-19 risk. Aspirational districts have a higher magnitude of transmission (susceptibility) as well as fatality (vulnerability). Discerning such locations can allow targeted resource allocation by governments to combat the next phase of this pandemic in India.

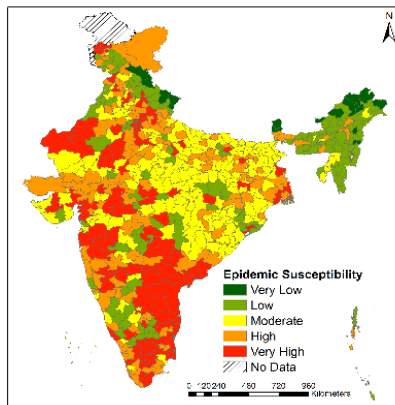
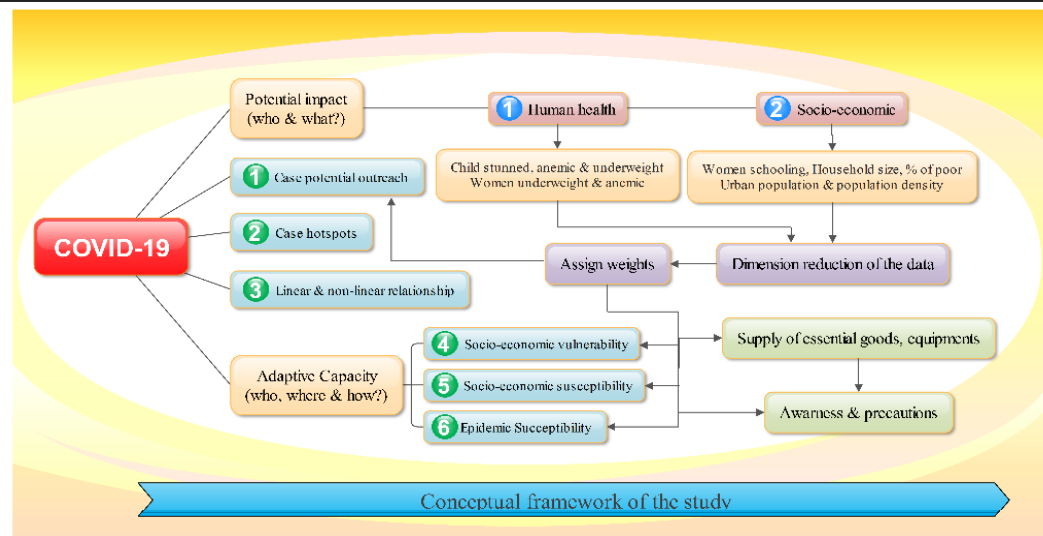
Keywords: Epidemic outbreak, contagion spread, socioeconomic analysis, epidemiological attributes, novel coronavirus, lockdown measures

Graphical Abstract

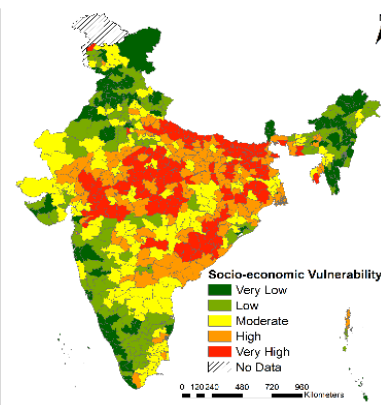
Questions:

1. How did COVID-19 spread across India?
2. Which districts can report the highest cases?
3. Which areas are likely to suffer the most deaths?
4. What factors affect the virus' incidence?
5. Can relatively safer areas and danger zones be identified?

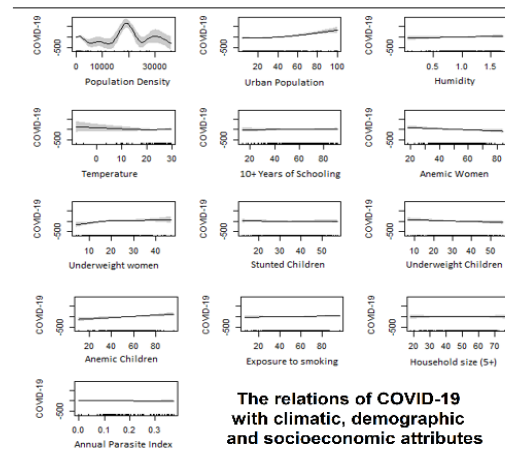
The Framework Adopted



Areas most likely to succumb to COVID-19



Areas most likely to have higher morbidity



The relations of COVID-19 with climatic, demographic and socioeconomic attributes

1. Introduction

At the time of writing, the COVID-19 global pandemic has raged for about five months, affecting 213 countries and territories worldwide and two international conveyances (<https://www.worldometers.info/coronavirus/>), in all afflicting over 8.4 million people and causing over 450,000 deaths till date. The emergence of this pandemic, as a previously unreported mystery illness or unknown strain of pneumonia in the Wuhan region of China (Columbus et al. 2020; Lupia et al. 2020; Torales et al. 2020), in December 2019 and its rapid dispersion across the globe in a matter of weeks (Bonilla-Aldana et al. 2020; Cruz et al. 2020), had caught countries and healthcare systems off-guard. The primary mode of spread from its centre of origin has been through air-travel to other major cities in different countries and thereafter from those places further into their hinterlands and adjacent regions, often on the coattails of returning tourists or businessmen, students or other travellers (Chakraborty and Maity 2020). With community transmission being set in motion often just days after the first cases were reported in some areas, the pandemic's burgeoning footprint overwhelmed healthcare services quite swiftly. The aged and those with existing health issues or compromised immune systems have been particularly vulnerable to this coronavirus (MacConnachie et al. 2007). Ameliorative measures have ranged from initially quarantining patients to progressively containing entire healthcare centres, city blocks and provinces (Harapan et al. 2020), as the virus' ambit grew beyond administrative, political and geographic boundaries. With transmissions occurring through contact between individuals, via respiratory droplets at close quarters or from contaminated surfaces (WHO 2020a), healthcare officials and governments have introduced and widely propagated the concept of 'social distancing' (Manderson and Levine 2020) to limit the spread of the virus, with cancellations of major sporting and cultural events (Munoz and Meyer 2020; Parnell et al. 2020) and diplomatic gatherings (Sharfuddin 2020), closure of religious institutions (Alyanak 2020) and the suspension of academic conferences and teaching activities (Gallo and Trompetto 2020). However, with the virus' spread having continued unabated and possibly being supplanted more alarmingly by transmissions from pre-symptomatic and asymptomatic individuals (WHO 2020b), further restrictive measures in the form of the partial and complete lockdown of regions or even entire countries were quickly adopted. Such lockdowns sought to heavily restrict the movement of those possibly carrying the contagion and to forbear them from transmitting it to others/other areas as well as curb the virus' spread by stopping healthy people from moving into spaces in which they were more likely to pick it up or come into contact with pre-symptomatic/asymptomatic individuals. Other nations, like India, where the outbreak assumed a potentially threatening status some weeks after its initial rampage in East Asia and Western Europe, were somewhat quicker to highlight social distancing and impose such lockdown measures (*The Lancet* 2020), seeking to take pre-emptive steps to combat the virus' spread and to try and keep affected populations to a minimum.

Yet, despite such restrictive measures being imposed, many countries, India included, have suffered increasingly rising outbreaks of the virus and experienced a sharp jump in cases due to existing gaps in their healthcare systems and apparent vulnerabilities in their socioeconomic and political-administrative setups (McAleer 2020), which exacerbate the risk of contamination and dampen recovery rates. While lives have definitely been saved as a result of the lockdown measures (Dwivedi 2020) and transmission rates are now checked in quite a few countries where the outbreaks occurred first (i.e. in China, South Korea, Italy, Spain,

Germany and France, which have all reported declining numbers of new cases and much-reduced transmission rates in the last few weeks), such effects are still not as apparent across parts of India even two months after the initial lockdown was implemented, with cases continuing to rise at increasing rates. This is also the scenario in many other affected regions where numbers continue to rise (i.e. Russia and Brazil) or has not yet tapered off significantly (i.e. the United States and the United Kingdom). On the other hand, the shutting down of offices and factories has already cast what is likely to be a lasting effect on the global economic landscape (Ajami 2020; Gong et al. 2020; McKibbin and Fernando 2020). In particular, nations that have large numbers of low-income migrant workers, who mostly live on-site at their workplaces or are heavily dependent on daily incomes/wages for sustenance, have perforce been uprooted. This has, in all probability, further reduced their ability to withstand a viral outbreak due to impaired health from malnutrition/hunger and the need to travel long distances to return to their original homes. Such a situation has especially emerged in India, where thousands of migrant workers mostly from the southern and western zones of the country, have thronged bus terminals and railway stations or have even attempted to travel across states on foot (Lal 2020; Singh 2020), raising the spectre of further widespread community transmission and the incursion of the virus into especially socioeconomically vulnerable areas (and more importantly into rural hinterlands) that are ill-equipped in terms of healthcare (both at the individual and community level for resident populations and returning migrants- Nacoti et al. 2020) and economic resources to deal with a massive surge in infections. Alongside the physiological threat to life and the economic threat to livelihoods posed by the current situation, arises an equally grim prospect of impaired mental health and similar attendant issues (Goyal et al. 2020; Torales et al. 2020; Wang et al. 2020). Continuous confinement at home can make daily life seem dysfunctional in an apparently dystopian society, with residents in constant fear of what may transpire (Ahorsu et al. 2020; Raffaeta 2020).

With no vaccine yet available, the prospect of a viral return in this coronavirus' original or mutated form casts further threat, more so on populations and areas that are most at risk from such an outbreak. We, therefore, seek to examine a number of aspects in this paper. *Firstly*, we delve into the pattern of COVID-19 outbreak in India, relating this with the influx of tourists from abroad (as the COVID-19 is an 'imported' virus- WHO 2020b) and those availing of air-travel in the country. *Secondly*, we track the contagion's spread through the nation during the successive/ongoing lockdown phases and map the geographic trajectories of COVID-19 infection in India together with its potential outreach zones, i.e. the areas where the virus was most likely to spread into when this analysis was undertaken. For doing this, we have had to perforce use slightly older data than that which reflects the current situation, due to the intervening lag time for computation, analysis and writing up of the results. Some research accounts have noted that the contagion's virulent nature is dampened and hindered by higher temperatures (Briz-Redon and Serrano-Aroca 2020; Prata et al. 2020; Wu et al. 2020; Xie and Zhu 2020) and that it reacts to raised humidity levels (Ahmadi et al. 2020; Luo et al. 2020; Qi et al. 2020), conditions which mostly persist across the India subcontinent. Along with this, clinical studies, as well as actual treatment accounts, have reported that hydroxychloroquine or chloroquine based drugs have achieved a measure of success in combating the COVID-19 virus (Colson et al. 2020; Gao et al. 2020; Singh et al. 2020). India has had a history of chloroquine use and vaccination due to its historical struggles with malaria (Das et

al. 2012; Anvikar et al. 2014) and it may thus be possible that her citizens are likely more resistant to the COVID-19 threat. Based on the above two aspects, we estimate the epidemiological susceptibility of a particular location (since this is, in essence, a medical/health issue first and foremost), identifying those areas where the disease is most likely to affect people. Subsequent to this, we examine the overall susceptibility of a region to a COVID-19 outbreak, factoring in the socioeconomic and health conditions of its residents, which makes them more prone to succumbing to the virus. Finally, we look at the overall socioeconomic vulnerability of the districts (i.e. where deaths are most likely) and its variation across the country.

As stated before, the lockdown has resulted in one of the largest mass migrations in contemporary India (in terms of movement being concentrated in such short timeframes) (Abi-Habib and Yasi 2020). It is obvious that some regions shall be more affected than others in this light, due to loss of income and livelihood and the dropping off in agricultural and industrial production. While most industries (both agro-based and manufacturing and on-site construction) and offices have gradually been opened up under the current lockdown phase operative in the county, the effects of their complete closure for over a month are sure to reverberate across regions, both urban and rural. We highlight those areas that are likely to be worst affected in such a situation. By combing our outputs from the epidemiological and overall susceptibility analyses together with that obtained from the ensuing vulnerability study, we thus identify regions that have already been the worst affected in terms of the virus' outbreak alongside those that are at the greatest risk from its further spread, firstly in medical terms, as well as in the light of further impaired economic activities, if the rising trend of the COVID-19 outbreak in India continues unabated over the coming weeks. As demonstrated already (e.g. Fauci et al. 2020; Liet al. 2020; Prem et al. 2020; Shi et al. 2020; Yang et al. 2020), such studies are crucial towards gaining a comprehensive idea of the pandemic's likely impact on a country (in health, economic and societal terms), not only in light of the current ongoing situation but to also gauge possible scenarios and ascertain target areas for pre-emptive resource allocation in the near future.

2. Susceptibility and vulnerability analyses in the context

Susceptibility analyses are important to ascertain the likely incidence of disease and the populations it can affect (Lohmueller et al. 2003). This also helps health-care providers perceive the risks involved for certain groups in order to better influence their health-related behaviour (BMJ 2020; Brewer et al. 2007). In relation to the COVID-19 virus, a range of susceptibility studies have been performed to understand how the physiological attributes of patients condition their response to the virus (Rothan and Byrareddy 2020; Shi et al. 2020; Zhao et al. 2020), and its links with the disease's spread in different areas (Fanelli and Piazza 2020; Karako et al. 2020). Other studies have compared habits like smoking (Wang et al. 2020) or ambient environmental conditions (Coccia 2020; Wu et al. 2020) with the virus' threat and examined how co-morbidities can occur due to underlying health issues (Li et al. 2020; Yang et al. 2020).

Examinations of the vulnerability of people and regions to this virus outbreak have followed diverse trends. The first group of studies have mostly focused on tracing the virus' actual footprint across different regions to identify communities that are most vulnerable to its onslaught and the ambient risks that they face (e.g. Boldog et al. 2020; Gilbert et al. 2020). The second category of attempts have employed model-based

analyses (e.g. Cassaro and Pires 2020; Cobey 2020) to predict estimates of the likely scenarios that can evolve, in order to pre-empt the virus' effect in a region (e.g. in China- Zhang et al. 2020; in Spain, France and Italy- Ceylan 2020 and Africa- Adekunle et al. 2020) and to gauge the outcomes of social distancing and restrictive travel measures (Chinazzi et al. 2020; Kraemer et al. 2020). The final category of vulnerability and COVID-19 impact studies have delved into how this virus' spread has affected healthcare services and shaped socioeconomic dynamics. Concerns about diminishing essential medical supplies (Requia et al. 2020; Rowan and Laffey 2020; Wang and Tang 2020) or impaired medical services to other patients with critical-care requirements (Soreide et al. 2020; Willan et al. 2020) have been paramount. The societal impacts of this virus have been undeniable (c.f. Ali and Alharbi 2020; Nicola et al. 2020), with people restricted indoors and almost all inter-personal contact shifting to the digital domain.

This global pandemic has also brought to the fore varied media channels and modes of information dissemination that encapsulate societal perceptions and reactions to the virus outbreak (Jimenez-Sotomayor et al. 2020), which can be fruitfully utilized to supplement official reports for tracking the contagion's spread (Oliver et al. 2020; Ting et al. 2020). These can enable measures to be put in place to reduce suffering in the areas most likely to be affected, through the use of advanced computational techniques and Big Data analysis (Cinelli et al. 2020; Srinivasa Rao and Vasquez 2020; Zhou et al 2020).

Attempts to estimate this coronavirus' spread have however been comparatively less in the overall Indian context (e.g. Biswas and Sen 2020; Dhanwant and Ramanathan 2020; Pandey et al. 2020), with generally regression models being employed to predict virus outbreaks across the country (Tomar and Gupta 2020; Singh and Adhikari 2020). Others have tried to estimate the COVID-19's prevalence within smaller regions (e.g. Simha et al. (2020) for the southern Indian state of Karnataka and Kumar (2020) for the western state of Maharashtra). Efforts have also been made to document ways to control this coronavirus' rampage across the country (Mandal et al. 2020) and the debilitating effects it can engender on the nation's overall healthcare system/setup (Chatterjee et al. 2020). The other aspect examined in the Indian context has been the efficacy of the implemented lockdown measures (e.g. Paital et al. 2020), following similar studies undertaken elsewhere (e.g. Ibarra-Vega 2020; Tobias 2020). Sardar et al. (2020) in examining the effects of the initial lockdown measures have opined that it would have little effect in the western Indian states that are the most affected by the virus but can prove to be somewhat beneficial in other areas, while Das et al. (2020) have provided guidelines on the critical community size required to make lockdowns effective.

However, it has been demonstrated that there exists a great deal of heterogeneity in how people and regions respond to and are affected by this pandemic (Sominsky et al. 2020), with an urgent need for further science-based assessments of the virus' spread (Bedford et al. 2019). With India being an extremely diverse country, in geographic as well as socio-demographic aspects, it is very much feasible that the virus' impact and spread will vary markedly across regions. Furthermore, due to the large variations in economic and healthcare attributes across the nation, the physiological response and case-fatalities will also differ significantly across areas. We, therefore, seek to examine the COVID-19's footprint across the nation and link it intricately with the situational realities in different locales, something which is sorely required but has received scarce attention in the current context.

3. Database and methods

3.1 Data collection

The entire analysis has been based on the district-level administrative zones (Figure 1) denoted in the Indian Population Census of 2011, which lists 640 districts nationwide, in order to keep parity with the Census datasets that were used. The number of district-level COVID-19 cases as reported by the Ministry of Health and Family Welfare (MoHFW), Government of India (GoI) was the principal secondary dataset used to analyze the geographic trajectories of this viral outbreak across India. This data was also used for the identification of COVID-19 hotspots and analysis of its possible outreach. On 24th March 2020, the Government of India (GOI) ordered a nationwide lockdown for 21 days (Phase-I), from 25 March - 14 April 2020, further extending this in Phase-II (15 April - 3 May) and then to Phase-III (May 3 to May 17). Therefore, when this study was commenced, district-wise numbers of COVID-19 cases were initially collected on four dates from the MoHFW, GoI, database to understand how this deadly virus is spreading across space and time and to feasibly represent each stage of the pre and ongoing lockdown phases. These dates were-a) Pre-lockdown Phase (on 23rd March); b) Early lockdown Phase (29th March); c) Mid-lockdown Phase (10th April) and d) Late lockdown Phase (18th April). Subsequently, the lockdown measures were extended to 31 May (Phase-IV, though there were significant differences during this period in terms of movement and operative industries/commercial establishments from the earlier more restrictive phases) and to show the situation at the end of the entire lockdown phase (from 23 March - 31 May 2020, with the period beyond this till 30 June being denoted as Unlock-1 as the nation reopens and many more measures are eased off), as an addendum we have mapped the nationwide situation on 1 June, 2020. This also allows us to comment on the overall efficacy of the lockdown measures implemented.

Two sets of indicators were used for the ensuing susceptibility and vulnerability analyses. For selecting these, we considered all the relevant information at the district level that could feasibly be associated with the COVID-19 outbreak (these datasets were collated from information made available by Office of the Registrar General of India (2011), MoHFW (2016), IIPS and ICF (2017) and similar other sources- see Table 1 and Table 2). The district-level susceptibility is related to viral transmission, while the vulnerability denotes the risk of fatality after infection. We identified the various potential indicators based on an extensive literature review (e.g. Coudhry and Avindandan 2020; Dowd et al. 2020; França et al. 2009; Rocklov and Sjodin 2020; Sajadi et al. 2020; Wang et al. 2020; WHO 2020a; Zhong et al. 2020) and the parameters finally selected for both the susceptibility and vulnerability analyses basically encapsulate the district-wise overall socioeconomic, demographic, climatic, health and hygiene conditions. Their descriptions and references are listed in Table 1 and Table 2, respectively, while the conceptual framework devised for the entire workflow is depicted in Figure 2.

3.2 Methods

3.2.1 Linking air travel and COVID-19 cases

Since the dispersion of the COVID-19 virus from its source region and its subsequent entry into every other geographic region was mostly via airports and returning travellers, the initial spreading centres and hotspots were most likely to develop around cities that handled large volumes of air traffic. The annual air passenger data for 2018-2019 was collected from the Airports Authority of India (<https://www.aai.aero/en>) for 54 international airports across the country. The number of COVID-19 cases within a district was used as an attribute of its centroid and a spatial buffer analysis was done to understand the relation between the distance from the airport and the number of COVID-19 cases. The Inverse Distance Weighting interpolation method was used to represent the ratio between the number of COVID-19 cases and the number of air passengers for each international airport, with higher ratios obviously indicating the greater likelihood of a large number of COVID-19 cases emanating from or being transmitted by passengers who have passed through that particular airport.

3.2.2 Geographic trajectories of the COVID-19 outbreak and hotspot identification

The district-level outbreaks of COVID-19 cases in the initial four consecutive time periods have been visualised through a series of maps and similarly, the changes in the virus hotspots and its potential outreach have also been identified for each time step. The Getis-Ord G_i^* statistic developed by Getis and Ord for analyzing evidence of spatial patterns (Getis and Ord 1992; Ord and Getis, 1995) was used to identify clusters of low and high index values and generate district-level hotspots and cold spots of COVID-19 cases. This tool is based on the z-score, p -value and confidence level bin (G_i _Bin) for each feature class. For the purpose of cluster identification, the neighbourhood distance was calculated using the Incremental Spatial Autocorrelation tool of ArcGIS 10.2 (Appendix 1).

3.2.3 Moran's I

The Moran's I (Moran, 1950) is useful for determining spatial autocorrelation by explaining whether or not certain attributes in a two-dimensional plane affect other attributes with the same characteristics (Drück et al., 2004). Moran's I was used to distinguish between a statistically significant COVID-19 cluster of high values surrounded by other similarly high-value districts (High-High: HH), a cluster of low values surrounded by low-value districts (Low-Low: LL), an outlier in which a high value was surrounded primarily by low values (High-Low: HL) and another outlier in which a low value was surrounded primarily by high values (Low-High: LH). The statistical significance was set at the 95 percent confidence level, with p -values smaller than 0.05. Moran's I was calculated for the same databases (of the four different time periods) as was used for the Getis-Ord G_i^* hotspots and cold spots analysis.

3.2.4 COVID-19 potential outreach analysis

The COVID-19's potential outreach was assessed using the vector based double and vector based integer assessment through Python coding. The Euclidean distance was computed from each district's centroid and the

number of COVID-19 cases for that district was used as its mass. As such, the higher the number of COVID-19 cases in a district, the higher would be its outreach potentiality towards the surrounding districts (Appendix 2).

3.2.5 Developing an area-based COVID-19 susceptibility and vulnerability index

As long as a vaccine for the COVID-19 has not been fashioned and distributed widely, susceptibility analysis of this viral transmission remains extremely important. As mentioned earlier, for the susceptibility analysis a number of possible crucial indicators were selected. Two susceptibility indices were developed to respectively gauge the epidemiological and socioeconomic susceptibility at the district-level. The epidemiological susceptibility analysis dataset includes the actual number of COVID-19 cases in the country along with other relevant socioeconomic, demographic and climatic factors. The socioeconomic susceptibility analysis dataset, on the other hand, excludes the variable denoting the number of COVID-19 cases in a district. We sought to undertake such a dual analysis in order to represent both, the actual situation which has emerged (as far as it was possible in the light of inherent data limitations and lags) as well as the situation that can arise if the country continues to remain under the threat of the virus (i.e. if all areas are eventually targeted or intruded into by the virus, which ones may succumb more easily). While constructing the epidemiological susceptibility index, the maximum weightage was given to the variable denoting the district-wise number of COVID-19 cases, as the virus' presence in a district endangers that area the most and makes it more liable to become a spreading centre than any other socioeconomic, demographic or climatic factor. For the socioeconomic susceptibility index, all variables were given almost equal weights (see Table 1). With the community transmission of the COVID-19 already underway and ongoing, it is tricky to conduct any susceptibility analysis using real-time information due to the obvious time-lag between information gathering and analysis. In this paper, we have used the then most recent available data for the districts as released by the MoHFW, GoI, when we commenced our analysis in the latter half of April 2020. Similar to the weightage pattern accorded for the socioeconomic susceptibility computation, the index for ascertaining the socioeconomic vulnerability was also developed by assigning equal weightage to all variables (Table 2) and using weighted mean method (Appendix 3).

3.2.6 Socioeconomic, demographic and climatic correlates of COVID-19 cases

The assessment of spatial heterogeneity in the relationship between two variables may be useful if the presence/numbers of COVID-19 cases is available for all districts. Due to the unavailability of this data, we have only focussed on eliciting linear and non-linear relationships of the district-wise number of COVID-19 cases with the ambient socioeconomic, demographic and climatic variables. In computing the linear relationship, we obtained Pearson's correlation plot using the '*corrplot*' package in R (Wei and Simko 2017). For the multivariate analysis, the non-linear relationship of COVID-19 cases and its correlates were computed using the Generalized Additive Model (GAM), via the '*gam*' package in R (Hastie 2019).

4. Results and Discussion

4.1 COVID-19 spread and its hotspots and clusters

COVID-19 cases were first diagnosed in India in late January, 2020 and crossed 60,000 cases by the end of the second week of May, 2020 (with present numbers in mid-June, 2020 being more than 3.4 lakhs). With this highly contagious disease transmitting primarily through international tourists and returning travellers from aboard, the initially affected sites were mostly cities that have international airports (or are regions adjoining such places) or are major tourist destinations. This is made apparent by the district-wise spread pattern of the COVID-19 virus (Figure 3). In the first phase (Figure 3A), cases were reported from western India (around Mumbai and Ahmedabad- two of the main commercial hubs of the country), from around New Delhi (the national capital) and Ladakh (popular tourist destination as well as a prominent base of the Indian Army) and from the southern states of Kerala (which has a high number residents who migrate/travel for work to the Gulf region), Tamil Nadu, Andhra Pradesh and Karnataka (all of which have major metropolitan centres that are important commercial hubs- Chennai, Hyderabad and Bengaluru). We saw a further intensification of the above pattern in the next time step on 29th March, 2020, with the adjoining regions of the above areas reporting substantial numbers of COVID-19 cases (Figure 3B). Broad swathes of the eastern part of the country still remained mostly unaffected, except for Kolkata (the major regional metropolitan centre) and its adjoining areas. In the next time step on 10th April, 2020, two large contiguous zones in northern India and western-central-south India reported high numbers of COVID-19 cases (Figure 3C). The northern entity manifested in a broad swathe from western Rajasthan, through Haryana, Delhi, western Uttar Pradesh and Punjab, up to Ladakh and parts of Jammu and Kashmir. The west-central-southern entity, while contiguous in nature had two broad trends- one portion of it stretched into the interior of the country from west to east across Maharashtra and Madhya Pradesh, the other followed the alignment of the eastern and western coastal plains, with one arm trending from Gujarat down to Kerala via Maharashtra, Goa and Karnataka's littoral districts and the other incorporating the coastal tracts of Andhra Pradesh and Tamil Nadu (with the adjoining parts of Telangana and Karnataka). Even at this comparatively advanced stage of the pandemic in India, the eastern portion of the country reported markedly lower numbers of cases, with only a few centres in Gangetic West Bengal, coastal Orissa and along the Ganga valley in eastern Uttar Pradesh and Bihar. The pandemic's footprint in the north-eastern states of India was by and large negligible. The final time step map of 18th April, 2020 (Figure 3D), showed the merging of the two entities described above, with infilling of their intervening districts (i.e. reporting of cases from areas that did not have any before) and a rise in the numbers of cases of those districts already afflicted previously. This resulted in a near-continuous stretch of the nation, from Kashmir to Kanyakumari, along the central and western corridors and down the eastern and western coastal belts, reporting COVID-19 cases. By now the country's eastern zone had noted a rise in cases (with its epicentre at Kolkata). This manifested as a narrow line of districts along the densely populated Ganga plains in Uttar Pradesh and Bihar, which merged into the larger/more contiguous zone further west.

The above spreading trends and initial incidence centres of the virus in India made it apparent that, as expected, the distance from an international airport and the numbers of COVID-19 cases were characterised by distance-decay or were inversely correlated to each other, i.e., with increasing distance from the airport the

number of COVID-19 cases decreased (Figure 4). During the successive lockdown phases, the percentage of cases within 50 km of an airport had decreased slightly but numbers rose gradually for locations within 50-100 km of an airport. Thus districts situated within this distance buffer (50-100 km) seemingly had a higher probability of being affected by COVID-19 and thereby swelling the overall numbers of cases. This demonstrates that the virus had spread out from its initial centres during the latter part of the lockdown and while it may have been partly contained in the primary places it had first afflicted, its footprint was becoming much more apparent across the hinterland. Of the 50 international airports examined, the ratio of air passengers to COVID-19 cases was far higher for the Mumbai, Delhi, Bengaluru and Hyderabad airports (Figure 5). Mumbai and Delhi handle the maximum international air traffic and passengers in the country and most international and domestic connecting flights are also routed through them. COVID-19 cases in India started to increase from the middle of March onwards, with this rise occurring after the virus' outbreak and spread in most European countries and the Middle and the Far East. In economic as well as geographic terms, Mumbai is intricately connected with the Middle East and also linked through trade and commerce with all major European markets and the Far East. Similarly, Hyderabad and Bengaluru, the main information technology hubs of India, are also entwined with the global market while New Delhi (the national capital) is connected by flights throughout the world for diplomatic, administrative and tourism purposes. Thus, the initial outbreak in India occurred mostly through these four airports. As mentioned before, many of Kerala's residents work in the Middle East or Gulf countries and this state is also a major tourist destination, with numerous flights routed to it through the Mumbai and Bengaluru airports. Therefore an initial hub of the COVID-19 incidence in India was also noticed therein.

The Getis-Ord G_i^* method was used to identify the spatial distribution of potential hotspots (statistically significant high values) and cold spots (low values) of the virus' spread/concentration through z-score, p -value and confidence level (G_i _Bin) analysis (Figure 6). The 23rd March, 2020 dataset indicated a spread of various COVID-19 clusters, with a z-score of 6.90 representing high clustering along the northern, western and southern regions of India (Figure 6A). A significant cold spot was visible in the eastern region, comprising of Assam, West Bengal, Bihar and their surrounding areas. The same analysis for the 29th March, 2020 dataset (Figure 6B) indicated clustering of cases in the west and south-western region of India, encompassing Kerala, Karnataka, Tamil Nadu, Maharashtra and Gujarat. The location of the cold spots remained almost similar to that discerned from the pre-lockdown phase data of 23rd March, 2020, with clustering mainly around West Bengal. The hotspot analysis of the mid-lockdown phase data of 10th April, 2020 (Figure 6C) indicated intense clustering of high values around Maharashtra and Gujarat. This finding is especially significant since these two states at present contain the highest numbers of total and active COVID-19 cases in India. Furthermore, this also validates to an extent the computations performed. However, the initial high values clustering around Kerala started diminishing in this mid-lockdown phase, possibly due to the pre-emptive actions regarding testing and quarantine taken by the Kerala State Government and the Union Government's policies. In fact, Kerala was on course to be the first Indian state to 'beat the curve', having shown at that time a continuous downtrend/decrease in its number of active COVID-19 cases since 6th April, around when numbers had peaked, to them reducing by 10th May to the low 20s (this situation has

subsequently changed markedly for the worse and the patterns/trends as of 1st June, 2020 are shown in Appendix 4 along with a discussion of the possible reasons for this sharp recent upturn in case numbers). The cold spot intensity had declined in the eastern part, almost throughout West Bengal. Analysis of the 18th April, 2020 dataset (Figure 6D) revealed a much more concentrated hotspot with a z-score of 7.84 in Maharashtra and Gujarat. By then, the entirety of Kerala and Andhra Pradesh had ceased to be part of any hotspot region. However, the area and intensity of the cold spot region located in the eastern part of the country had also declined significantly, possibly indicating that the next phase of the virus' spread could possibly target this zone. Again, subsequent events and updated news reports about the COVID-19 spread have validated just this, with the eastern region around Kolkata presently emerging as one of the most affected by the pandemic (see Appendix 4).

The investigation of global clustering patterns across regions is very important in spatial data analysis and the Moran's I is a widely used spatial statistic for this purpose (Jackson et al. 2010). While the Getis-Ord G_i^* analysis of the district-level COVID-19 cases identified the clustering of low and high index values, it only delineated large clusters through neighbourhood analysis and ignored those districts that had high numbers of cases but were surrounded by low-value neighbours or vice versa. Thus, the Moran's I was computed to obtain additional insights into the statistically significant High-High, High-Low, Low-High and Low-Low clusters derived previously for each of the four datasets (Figure 7). Results reveal that the Moran's I generally follows the pattern of the Getis-Ord G_i^* hotspot and cold spot classification but that it also provides further visual identification of all High-Low and Low-High value neighbours among the COVID-19 affected districts. The Moran's I map for 18th April, 2020 (Figure 7D) indicates some locations in cold spot regions that have attribute values higher than the overall average while their neighbours' attribute values are lower than the average. Such areas include Nalanda and Siwan districts of Bihar, Korba district in Chhattisgarh, Khordha district in Orissa and Lucknow in Uttar Pradesh. Contrastingly, the hotspot regions derived from the same data show that the attribute values of many districts are lower than the average while their respective neighbours have higher than average attribute values. All the High-Low designated districts in the cold spot areas are quite important tourist destinations and the Low-High designated districts in the hotspot zones are on the whole not as developed or as urban as their higher value neighbours. However, more in-depth research is required to fully discern the observed pattern of High-Low and Low-High Moran's I values and its temporal changes across the four examined datasets.

Each of the above three attributes, i.e. the spread of COVID-19 cases, the clustering of cases and the hot/cold spots have evolved further since this study was commenced with the initial database of the first few lockdown phases. Appendix 4 has thus been added to show the changed conditions as on 1st June, 2020, in the above aspects, thereby denoting the complete scenario during the pre and entirety of the lockdown phases, that effectively ended on 31st May, 2020.

4.2 Estimating the potential outreach of the COVID-19 pandemic across India

The potential gravity of the COVID-19 spread was assessed for the four time stamps to ascertain its likely outreach areas (Figure 8). The possible potential outreach on 23rd March, 2020 (Figure 8A) was mostly around the Delhi NCR (National Capital Region). Some small outreach pockets were seen in Kerala and in Maharashtra (e.g. Mumbai and Pune). As per the 29th March, 2020 dataset (Figure 8B), potential outreach epicentres were found in three significant locations- Delhi-centric, Mumbai-centric and Kerala-centric (due to reasons outlined in Section 4.1). In the third phase of the analysis (i.e. dataset of 10th April, 2020), Kerala displayed a marked improvement in controlling the outbreak, as discussed previously (Figure 8C). However, rapid growth was observed for Mumbai and its surrounding areas and this scenario continued into the next phase as well (Figure 8D). We thus predicted that based on the potential outreach analysis, the gravity of the COVID-19 outbreak was likely to be very high in western India, especially in Maharashtra and Gujarat. The current situation in June, 2020, fully bears out the above analysis, as these two states rank the highest in the country in COVID-19 related deaths (India Covid-19 Tracker at <https://www.covid19india.org/>). Maharashtra is overwhelmingly afflicted in Mumbai and its nearby areas of Thane and Pune, with its total number of cases being almost three times that of second-placed Gujarat (which has been hardest hit in Ahmedabad, Surat and Vadodara). As per the performed potential outreach analysis, the North-eastern states and the Himalayan states are relatively safe from the COVID-19 outbreak. Total cases in Sikkim, Arunachal Pradesh and Meghalaya in India's northeast are in double digits only, while numbers in most other states in the region are relatively quite low. However, the eastern districts of Bihar, Jharkhand and West Bengal are seen to gradually come under the pandemic's grasp, over the different times-steps, pointing towards this zone possibly carrying forward the rising patient numbers in the near future, as the spread in western India may start to peak or gradually decline. Most of those in eastern India who are afflicted by COVID-19 have had a travel history from either Mumbai, Kerala or Delhi. Initially, only a few cases were observed in eastern India due to the lower volume of international air traffic in this region (as denoted by the lower air passenger to COVID-19 case ratio) but more recently, the numbers of cases have risen herein in the late lockdown phase due to inter-state in-migration from the western and southern parts of the country, particularly as a result of the returning migrant workers (Mullick 2020).

Overall, the most important results derived from the potential outreach analyses were that the epicentres of COVID-19 have so far been mostly concentrated in and around Mumbai in Maharashtra. As Kerala and also Karnataka recovered quite quickly, the earlier epicentre in southern India was quite obliterated during the latter lockdown phases (however, infections have risen sharply in Tamil Nadu towards the end of the final lockdown phases in May 2020 as restrictions have relaxed- see Appendix 4). The eastern part of the country now faces the imminent threat of becoming an epicentre if COVID-19 cases are transferred into it due to return migration, especially by the attempted/ongoing homecoming of the many currently unemployed migrant labourers in the western and southern reaches of the country (which has manifested subsequently, along with increasing case numbers in areas earlier showing a declining trend, i.e. in Kerala). The comparative outlooks of the undertaken hotspot and the potential outreach analyses differed slightly in perspective. While the hotspot enumeration discerned COVID-19 affected zones based on the actual ground situation/data as

ascertained for that point in time, the potential outreach maps denoted the gravity factors for the four different time periods and depicted how conditions may worsen in the major epicentres or has eased off in other areas. The outreach analysis thus casts some light on the outbreak's potential across the country and denotes zones likely to be affected in the near future, with higher potential districts obviously facing a bigger threat from COVID-19.

4.3 Interpreting COVID-19 susceptibility and vulnerability indices

Using the indicators listed in Table 1 and Table 2, we developed an area-based composite COVID-19 susceptibility and vulnerability indices at the district level-for India, with a view towards providing policy makers with some indication on which districts are likely to be most susceptible or vulnerable to a COVID-19 outbreak and specifically where should the Government target its resources and accordingly plan a data-driven intervention strategy. The elicited results from these indices are presented below.

4.3.1 Epidemic susceptibility index

A five-pronged classification scheme (ranging from Very High to Very Low) has been used to visualise the derived district-level epidemic susceptibility values (derived using quintile class- Table 3). There are significant regional clusters in the northern, eastern, southern and parts of the north-eastern states of India (Figure 9). Large swathes of south-western and northern India, covering most districts of Kerala, Tamil Nadu, Telangana, Andhra Pradesh, Karnataka, Maharashtra, Gujarat, Rajasthan, Delhi, Haryana and Punjab are highly susceptible to this pandemic. Primarily, these were the areas where the initial outbreak occurred and the subsequent transmission of COVID-19 has been phenomenal, with people therein being seemingly less able to cope with the COVID-19 virus (i.e. their physiology is more easily affected by it). Possibly, the prevalence of urbanization (which creates congestion), the burden of non-communicable diseases (hyper tension, diabetes and obesity) and a greater proportion of the elderly population in these regions may have heightened the overall epidemic susceptibility. Most of the highly urbanized and economically well-off states report higher COVID-19 susceptibility, since the initial epicentres of the virus outbreak were in their large cities and significant economic ramifications may arise as such areas are likely to remain under stringent lockdown measures for relatively long periods. Contrarily, almost all districts in the seven north-eastern states of Assam, Meghalaya, Manipur, Tripura, Mizoram, Arunachal Pradesh and Sikkim fall under the low to very low COVID-19 susceptible class. The northern, central and eastern parts of India comprising of Odisha (Orissa), Chhattisgarh, Jharkhand, Madhya Pradesh, Uttar Pradesh, Bihar and West Bengal are moderately susceptible to the pandemic. It is interesting that despite being economically poorer and having low indicators with respect to almost every socioeconomic and health-related parameter than the rest of India, these areas are less susceptible to the pandemic, suggesting that the COVID-19 outbreak might be an erratic phenomenon that cannot be explained solely by traditional socioeconomic theories and which might require further investigation. As the districts in the moderate susceptibility to COVID-19 class lie in many economically poorer zones, this can have important connotations in terms of impaired health-care services and facilities, which can thereby significantly deteriorate immune response and patient recovery. Furthermore, as discerned from the previous

section, it is this eastern part of the country that is likely to become a secondary epicentre of the virus in the country (following the trends displayed by the hotspot, Moran's I and potential outreach analyses) subsequent to the primary one in the nation's west. Thus, even though this zone is at present in the moderate susceptibility category, its lower socioeconomic standing and existing poorer healthcare assets and availability/accessibility accord it a high priority for pre-emptive future resource allocation in order that the likely forthcoming challenges can be met adequately.

4.3.2 COVID-19 related district-wise socioeconomic susceptibility

Figure 10 visualizes the socio-economic susceptibility index for India's districts, following a similar classification approach to that adopted in Figure 9 (see Table 3). The underlying assumption of this particular index is that if all districts are evenly infected by the COVID-19 outbreak, then what would be the magnitude or susceptibility of further transmission in a certain area, with this being dependant on its socioeconomic status. Results reveal that large portions of eastern, north and north-western India are high to very highly socioeconomically susceptible to this deadly virus and are likely to be most affected by it. These areas are particularly characterized by high population densities, chronic malnutrition, poor health infrastructure, larger family sizes, poor hygiene practices, poverty concentration and marginalization including lower health-related knowledge and awareness, thereby precipitating such outcomes. On the contrary, most districts in the north-eastern, southern and extreme northern region (e.g. in the northern part of Rajasthan and in Punjab, Himachal Pradesh, Jammu and Kashmir and Uttarakhand), and are found to be socioeconomically less susceptible in terms of COVID-19 transmission. Interestingly, some areas that initially had higher numbers of COVID-19 cases have emerged in the low susceptible category (e.g. Kerala) and this merits further explanation. In Kerala, after the initial outbreak was reported, effective State Government measures at all levels towards curbing further transmissions by facilitating mass testing, awareness creation at the community level and applying stringent physical distancing and lockdown norms, had proven (at that time) to be very effective (Biswas 2020). Alongside this, most importantly, there had possibly been efficient management of both international and inter-state migrants in the state. However, conditions have worsened as restrictions have eased off, and large numbers of expats have returned from the Gulf region and from other parts of the country (Nidheesh 2020). These are lessons that can be picked up and implemented by other similarly afflicted areas or those likely to be threatened by the COVID-19 virus in the foreseeable future.

Aspirational Districts (ADs) are more susceptible to the COVID-19 due to their already limited coping-up capacities as they are some of the socio-economically poorest/most backward regions of India. The overall health infrastructure and particularly the numebtrs of primary and community health centres in such locales is also less, with limited staff and bare minimum facilities further compounding the issue. However, the incidence of COVID-19 cases in these areas was initially quite low, at less than 2% (a total of 610 cases as of 6th May, 2020) of the total cases nationwide. Of the 112 ADs affected by the virus, the worst-hit were Ranchi (55 cases), Baramulla (62), YSR (55), Nuh (57), Jaisalmer (34) and Kupwara (47), which are all located in the Red Zone (i.e. the locales facing the most stringent lockdown norms) as per the Indian Government's classification. Considering their susceptibilities, the Government of India (through the NITI Aayog) has taken

steps to ensure appropriate and timely action for resolving supply shortages in test kits, Personal Protective Equipment (PPE) and in providing masks to the respective empowered populations/groups in these districts. In our present study, we have also been able to identify these districts based on their respective susceptibility indicators, e.g. percentage of poor population, household size, impaired child health (i.e. suffering from anaemia, underweight or stunted development), and poor female health (anaemia and underweight). Apart from these ADs, we have also identified some other districts that are also quite susceptible due to their socioeconomic background and these areas, therefore, need a similar kind of support to that being provided to the ADs (i.e. urgent supply of testing kits, PPE and masks). Most districts of Madhya Pradesh, Orissa, Andhra Pradesh, Bihar, Jharkhand, West Bengal and Uttar Pradesh fall under the above category.

4.3.3 COVID-19 related district-wise socio-economic vulnerability

Based on the selected set of indicators, the discerned socioeconomic vulnerability to the COVID-19's impact (Figure 11) is likely to be higher in many districts of Madhya Pradesh, Orissa, Telangana, Andhra Pradesh, Bihar, Jharkhand, West Bengal and Uttar Pradesh and these areas are likely to report higher fatalities. There is a 30% overlap of the places most at risk in these states with those demarcated under the Union Government's Aspirational Districts Programme, i.e. 75 of the discerned 255 highly to very highly vulnerable districts are also ADs. The scarcity of healthcare facilities and personnel in these districts remains a major issue and such areas also have a large number of inter and intra state migrants, who have steadily returned home during the lockdown. Therefore, it is quite possible that cases of COVID-19 infection shall sharply rise in these locations and that the poorer medical infrastructure shall increase the infection rate and incidence of deaths, if adequate measures are not taken. Contrastingly, this impact shall be markedly lower in the states of Kerala, Punjab, Haryana, Himachal Pradesh and in most of the North-eastern states and the Rann of Kachch (in Gujarat). Some districts of Tamil Nadu, Gujarat, Rajasthan and Maharashtra are moderately vulnerable. The enumerated index values are higher in some districts due to these housing a greater proportion of those that are more vulnerable to this contagion (i.e. more number of children and women who are anaemic and/or underweight, lower female education levels and a higher prevalence of diabetic patients). Therefore, these districts are likely to find it difficult to cope with the COVID-19 threat and its related morbidity aspect. On the other hand, the lesser vulnerable districts seemingly do have sufficient capability to deal with the threat, as observed in the districts of Kerala, from where very few deaths have been reported, compared to the other parts of India, despite this state being an epicentre initially.

4.4 COVID-19 incidence and its correlations with climatic, demographic and socioeconomic aspects

Figure 12 shows the correlations derived between the COVID-19 cases (i.e. the number of cases per district) and its socioeconomic, demographic and climatic factors. This pandemic is positively associated with the percentage of urban population and the population density of a district. It is also positively associated with the percentage of women having attained 10th or lower standard of schooling. Furthermore, districts with a higher share of poorer households (40%) had lower viral transmissions. In the second half of April 2020, although the virus had spread across 377 districts, the gravity of the outbreak was mostly concentrated in districts that

were major urban agglomerations, such as Mumbai, Delhi and Hyderabad. People in urban areas are relatively more mobile within the city than ruralites in villages. They may also be forced to maintain lower residential and social distances due to the higher densities of built-up zones as well as population, especially within slums areas where housing shortages are quite severe (Haque et al. 2020). These factors together with the inherent economic deprivation faced by slum residents would likely enable a more widespread outbreak in such congested locales (Ahmed et al. 2020), as has been evidenced by the large numbers of cases reported from the Dharavi slum area of Mumbai, which is one of the largest such entities in Asia. This has also likely been the cause for the extremely high number of infections being subsequently reported from Delhi and its surrounding urban agglomeration and from Chennai and its neighbourhoods. As urbanites generally attain a higher level of education, the pandemic's outbreak is also positively associated with the educational status, more so because this 'imported' virus mainly came into the country via air-travel (i.e. from the movement of economically well-off sections of the population, who can be expected to have also attained higher education levels).

The above bivariate linear relationship does not provide the non-linear associations between variables. For this, the multivariate non-linear association was evaluated by applying the Generalized Additive Model (GAM), whose results are relatively easy to interpret. The obtained findings predict a flexible relationship between the enumerated variables (Figure 13). Districts with a population density between 1500-2500 people/km² specifically have a higher risk of this virus spreading amidst them than other districts. Of the 377 districts affected by the virus (as per the last date on which the data for this paper was collected), apart from a few situated in urban metropolitan cities, all others contain medium to large towns which have a moderate level of population density. The highly urban districts face a greater threat of a COVID-19 outbreak, particularly when the share of urban population crosses 60%. In the urban agglomerations of India, a significant share (17.4%) of the urban population lives in slums (e.g. about half the residents of Greater Mumbai live in such locales). While they play a vital role in the functioning of the city, these slum areas are poorly planned and obviously very densely populated as a result of the ever-accelerating urbanization trend in India (Haque and Patel 2018). Such sites thus become potential hotspots for infectious diseases like the COVID-19. This unplanned urban growth also poses a considerable risk in terms of impaired preparedness and ready response to any infectious disease outbreak, with mass quarantining at such close quarters while maintaining social distancing, being almost impossible. Hence, the responses of healthcare officials, governments and communities to this ongoing pandemic can be devised towards generating a paradigm shift in how urban spaces and residences are planned and designed, with possible ramifications for future peri-urban transformations, inner-city renewal and slum rehabilitation (Jha 2020; Regmi 2020; Van den Berg 2020).

The findings also suggest that although districts with a higher prevalence of children suffering from anaemia and/or being underweight face an elevated risk of this viral transmission, the relationship slope is not very significant. Thus these health indicators are not strongly correlated with COVID-19 cases in the Indian context. A possible reason for this seemingly discordant finding is that maybe the Government of India and/or State Governments had till the last date of initial data collection for this paper (i.e. 18th April, 2020) not conducted an adequate number/proportion of population-level tests, rather testing only those who had already

developed COVID-19 related symptoms. Therefore, a similar analysis performed on a bigger and possibly more representative sample obtained via large-scale population-level testing may yield better explanations.

Several studies have found that climatic parameters such as mean annual temperature and relative humidity have a crucial role in spreading the COVID-19 virus (Sajadi et al. 2020; Wang et al. 2020), with evidence suggesting that a mean temperature of 10°C and a medium to high (60% to 90%) relative humidity range is most suitable for its transmission. The greater majority of India's districts have a higher daily mean temperature than 10°C in April and except for a few districts in the Himalayan and coastal belts, all districts have an average relative humidity less than the suitable level. Therefore, unlike as discerned in other studies, we could not find any significant relationship between the currently prevalent climatic parameters and this viral outbreak.

Summary and Conclusion

The COVID-19 crisis was slow to onset, before rampaging worldwide and its impact on society and the global economy has been profound and likely to be long-lasting. The primary reason for this was that governments either underestimated its potential threat or were caught unawares or did not have robust enough socio-political systems and infrastructure to combat the virus' transmission. Instead, while helpful, the uniform global response of lockdown measures has created extreme economic vulnerability, and developing nations are some of the worst affected, having to contain the pandemic while simultaneously addressing rising unemployment and the livelihoods at stake of an economically vulnerable population. Thus, it is paramount that in the midst of such a crisis, correct and sustainable economic, social and political decisions be made. To enable these predictions of the pandemic's path across a nation with pinpointing of the areas that it can affect the most, are pertinent, for prudent and targeted resource allocation.

Our analyses have highlighted the initial centres of the COVID-19 pandemic in India, how its incidence occurred, its spread over time and concentrations. By identifying hotspots and significant clusters of the virus' outbreak, we pinpointed the areas where possible community transmission have occurred and which need targeted measures to control the situation. Our computations and predictions were based on the data garnered about the cases occurring in the initial and middle stages of the initial lockdown phases that were underway in India. However, subsequent reports (and datasets- see Appendix 4) of the pandemic's spread largely validate our estimation of the areas that were most likely to be affected in the forthcoming days (i.e. the case of the virus gaining a foothold and then spreading in various clusters of Eastern India, around Kolkata). Our analysis thus lays the groundwork for identifying future hot spot zones so that communities can prepare in advance and the local governments can anticipate requirements and stock/allocate critical resources accordingly. Mass testing of COVID-19 in such hotspots can curb the disease's spread into adjoining areas. By also designating areas that are relatively safer and less susceptible and vulnerable to the pandemic and where it is unlikely to spread into (i.e. cold spots), we also provide a framework for earmarking those locales where economic activities can be jumpstarted again more robustly than already underway, to boost flagging economies.

We have also identified districts that are more vulnerable to the virus. These locations can carefully restart agricultural and industrial activities away from large habitations and strive to build food stocks to avoid any food-security related issue in case sudden lockdowns are again required, taking care to note that the large numbers of returning migrant workers are mostly ruralites and infections among them can quickly alter the safety status of an area. The less vulnerable locales can engage in industrial activity in isolation from settlements to supply necessary commodities and various agricultural products like pesticides and fertilizers. In the most vulnerable cities and the suburban areas located within 50-100 km around them, crowds in market places must be stringently restricted, with the initiation of door-to-door supply of necessary commodities by local distributors, if required. Districts that have high epidemiological susceptibility need enhanced vigilance (possibly through specially constituted rapid-action teams) to avoid transmissions. Areas that are socioeconomically susceptible face the greatest threat as they house substantial numbers of people who may not be physiologically capable of coping with the COVID-19 virus. Further extension of lockdown in these areas may be a prerequisite to enhancing their health outcomes. Overall, our demarcated hot and cold spot areas and those that are epidemiologically susceptible or socioeconomically vulnerable tallies quite well with the red-orange-green containment zones recently delineated by the Government of India, providing further validation of the performed analyses (NDMA 2020).

There are still some swathes of India where the COVID-19 has not been able to intrude significantly, especially in physiographically rugged and isolated areas and in many rural hinterlands. What is paramount now, is that these locations be protected in order that community transmissions do not occur while the areas already under threat are dealt with. While the quickly enforced lockdown has no doubt caused much misery, especially for stranded migrant workers far from home, it may have also given India a little breathing space to anticipate the next move/spread of the pandemic and combat it accordingly. We hope that our study has created a framework that enables this to be accomplished fruitfully.

Acknowledgements

The authors are grateful to their respective institutions for providing the infrastructure required to conduct this research. The assistance of Ms. Sayoni Mondal, Research Scholar at the Department of Geography, Presidency University, Kolkata, India in arranging the article references is gratefully acknowledged.

References

- Abi-Habib M, Yasir S (2020) India's Coronavirus Lockdown Leaves Vast Numbers Stranded and Hungry. The New York Times 29th March, 2020. Available at <https://www.nytimes.com/2020/03/29/world/asia/coronavirus-india-migrants.html> (accessed on 11th May 2020)
- Adekunle IA, Onanuga A, Wahab O et al. (2020) Modelling spatial variations of coronavirus disease (COVID-19) in Africa. Science of the Total Environment 729:138998. <https://doi.org/10.1016/j.scitotenv.2020.138998>

- Ahmadi M, Sharifi A, Dorosti S, Ghouschi SJ, Ghanbari N (2020) Investigation of effective climatology parameters on COVID-19 outbreak in Iran. *Science of the Total Environment* 729:138705 <https://doi.org/10.1016/j.scitotenv.2020.138705>
- Ahmed F, Ahmed N, Pissarides C, Stiglitz J (2020) Why inequality could spread COVID-19. *The Lancet* 5(5):e240. [https://doi.org/10.1016/S2468-2667\(20\)30085-2](https://doi.org/10.1016/S2468-2667(20)30085-2)
- Ahorsu DK, Lin C-Y, Imani V, Saffari M, Griffiths MD, Pakpour AH (2020) The fear of COVID-19 scale: development and initial validation. *International Journal of Mental Health and Addiction* <https://doi.org/10.1007/s11469-020-00270-8>
- Ajami R (2020) Globalization, the Challenge of COVID-19 and Oil Price Uncertainty. *Journal of Asia-Pacific Business*. <https://doi.org/10.1080/10599231.2020.1745046>
- Ali I, Alharbi OML (2020) COVID-19: Disease, management, treatment and social impact. *Science of the Total Environment* 728:138861. <https://doi.org/10.1016/j.scitotenv.2020.138861>
- Alyanak O (2020) Faith, politics and the COVID-19 pandemic: The Turkish Response. *Medical Anthropology*. <https://doi.org/10.1080/01459740.2020.1745482>
- Anvikar AR, Arora U, Sonal GS, Mishra N, Shahi B, Savargaonkar D, Kumar N, Shah NK, Valecha N (2014) Antimalarial drug policy in India: Past, present and future. *Indian Journal of Medical Research* 139(2):205-215. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4001331/>
- Battle, K., E., et al. (2015). Global database of *Plasmodium falciparum* and *P. vivax* incidence records from 1985–2013. *Sci. Data* 2, 150012. <https://doi.org/10.1038/sdata.2015.12>
- Bedford J, Farrar J, Ihekweazu C, Kang G, Koopmans M, Nkengasong J (2019) A new twenty-first century science for effective epidemic response. *Nature* 575: 130-136. <https://doi.org/10.1038/s41586-019-1717-y>
- Biswas K, Sen P (2020) Space-time dependence of corona virus (COVID-19) outbreak. *Physics and Society* (preprint) arXiv:2003.03149 [physics.soc-ph]
- Biswas S (2020) Coronavirus: How India's Kerala state 'flattened the curve'. BBC News, 16th April. Available at <https://www.bbc.com/news/world-asia-india-52283748> (accessed on 14th June 2020)
- BMJ (The British Medical Journal) (2020) Covid-19: risk factors for severe disease and death. *BMJ* 368:m1198 <https://doi.org/10.1136/bmj.m1198>
- Boldog P, Tekeli T, Vizi Z, Denes A, Bartha FA, Rost G (2020) Risk Assessment of Novel Coronavirus COVID-19 Outbreaks Outside China. *Journal of Clinical Medicine* 9(2):571 <https://doi.org/10.3390/jcm9020571>
- Bonilla-Aldana DK, Holguin-Rivera Y, Cortes-Bonilla I, et al. (2020) Coronavirus infections reported by ProMED, February 2000- January 2020. *Travel medicine and Infectious Disease* 101575. <https://doi.org/10.1016/j.tmaid.2020.101575>
- Brewer NT, Chapman GB, Gibbons FX, Gerrard M, McCaul KD (2007) Meta-analysis of the relationship between risk perception and health behavior: the example of vaccination. *Health Psychology* 26(2):136-145. DOI: 10.1037/0278-6133.26.2.136

- Briz-Redon A, Serrano-Aroca A (2020) A spatio-temporal analysis for exploring the effect of temperature on COVID-19 early evolution in Spain. *Science of the Total Environment* 728:138811. <https://doi.org/10.1016/j.scitotenv.2020.138811>
- Cassaro FAM, Pires LF (2020) Can we predict the occurrence of COVID-19 cases? Considerations using a simple model of growth. *Science of the Total Environment* 728:138834. <https://doi.org/10.1016/j.scitotenv.2020.138834>
- Ceylan Z (2020) Estimation of COVID-19 prevalence in Italy, Spain, and France. *Science of the Total Environment* 729:138817. <https://doi.org/10.1016/j.scitotenv.2020.138817>
- Chakraborty I, Maity P (2020) COVID-19 outbreak: Migration, effects on society, global environment and prevention. *Science of the Total Environment* 728:138882. <https://doi.org/10.1016/j.scitotenv.2020.138882>
- Chatterjee K, Chatterjee K, Kumar A, Shankar S (2020) Healthcare impact of COVID-19 epidemic in India: A stochastic mathematical model. *Medical Journal Armed Forces India* (in press) <https://doi.org/10.1016/j.mjafi.2020.03.022>
- Chinazzi M, Davis JT, Ajelli M et al. (2020) The effect of travel restrictions on the spread of the 2019 novel coronavirus (COVID-19) outbreak. *Science* 368:395-400. DOI: 10.1126/science.aba9757
- Cinelli M, Quattrocioni W, Galeazzi A, Valensise CM, Brugnoli E, AL Schmidt, Zola P, Zollo F, Scala A (2020) The COVID-19 social media infodemic. arXiv:2003.05004v1 [cs.SI]
- Cobey S (2020) Modeling infectious disease dynamics. *Science*. eabb5659 DOI: 10.1126/science.abb5659
- Coccia M (2020) Factors determining the diffusion of COVID-19 and suggested strategy to prevent future accelerated viral infectivity similar to COVID. *Science of the Total Environment* 729:138474. <https://doi.org/10.1016/j.scitotenv.2020.138474>
- Colson P, Rolain JM, Raoult D (2020) Chloroquine for the 2019 novel Coronavirus SARS-CoV-2. *International Journal of Antimicrobial Agents* 55:105923. <https://doi.org/10.1016/j.ijantimicag.2020.105923>
- Columbus C, Brust KB, Arroliga AC (2020) 2019 novel Coronavirus: an emerging global threat. In: *Proceedings of the Baylor University Medical Center*, pp 209-212. <https://doi.org/10.1080/08998280.2020.1731272>
- Coudhry V, Avindandan V (2020) Why COVID-19 Outbreak in India's Slums Will Be Disastrous for The Urban Poor. <https://www.outlookindia.com/website/story/opinion-covid-19-outbreak-in-indias-slums-will-be-disastrous-for-the-urban-poor/350335> (accessed on 8th April 2020)
- Cruz MP, Santos E, Velazquez-Cervantes MA, et al. (2020) COVID-19, a worldwide public health emergency. *Spanish Clinical Magazine*. <https://doi.org/10.1016/j.rce.2020.03.001>
- Das A, Anvikar AR, Cator LJ et al. (2012) Malaria in India: The Centre for the Study of Complex Malaria in India. *ActaTropica* 121(3):267-273. 10.1016/j.actatropica.2011.11.008
- Das S, Ghosh P, Sen B, Mukhopadhyay I (2020) Critical community size for COVID-19- a model based approach to provide a rationale behind the lockdown. *Populations and Evolution* (preprint) <https://arxiv.org/abs/2004.03126> [q-bio.PE]

- Dhanwant JN, Ramanathan V (2020) Forecasting COVID 19 growth in India using Susceptible-Infected-Recovered (S.I.R) model. Populations and Evolution (preprint)arXiv:2004.00696 [q-bio.PE]
- Dowd JB, Rotondi V, Adriano L, Brazel DM, Block P, Ding X, Liu Y, Mills MC (2020) Demographic science aids in understanding the spread and fatality rates of COVID-19. medRxiv (preprint) <https://doi.org/10.1101/2020.03.15.20036293>
- Drück, S., Carvalho, M.S., Câmara, G., Monteiro, A.V.M. (2004). *AnáliseEspacial de Dados Geográficos*. Brasília: EMBRAPA.
- Dwivedi LK, Rai B, Shukla A, Dey T, Ram U, Sekhar C, Dhillon P, Yadav S, Unisa S (2020) Assessing the impact of complete lockdown on COVID-19 infections in India and its burden on public health facilities: A situational analysis paper for policy makers. IIPS, Mumbai. Available at https://iipsindia.ac.in/sites/default/files/iips_covid19_aicl.pdf (accessed 11th May 2020)
- Fanelli D, Piazza F (2020) Analysis and forecast of COVID-19 spreading in China, Italy and France. *Chaos, Solitons & Fractals* 134: 109761 <https://doi.org/10.1016/j.chaos.2020.109761>
- Fauci AS, Clifford Lane H, Redfield RR (2020) COVID-19- navigating the uncharted. *The New England Journal of Medicine* 382:1268-1269. DOI: 10.1056/NEJMe2002387
- França TGD, Ishikawa LLW, Zorzella-Pezavento SFG, Chiuso-Minicucci F, da Cunha MLRS, Sartori A (2009) Impact of malnutrition on immunity and infection. *Journal of Venomous Animals and Toxins including Tropical Diseases* 15(3):374-390. <http://dx.doi.org/10.1590/S1678-91992009000300003>
- Gallo G, Trompetto M (2020) The Effects of COVID-19 on Academic Activities and Surgical Education in Italy. *Journal of Investigative Surgery*. <https://doi.org/10.1080/08941939.2020.1748147>
- Gao J, Tian Z, Yang X (2020) Breakthrough: Chloroquine phosphate has shown apparent efficacy in treatment of COVID-19 associated pneumonia in clinical studies. *Bioscience Trends* 14(1):72-73. 10.5582/bst.2020.01047
- Getis, A., and Ord, J. K. The analysis of spatial association by use of distance statistics. *Geographic Analysis*, Vol. 24, No. 3, 1992, pp. 189-206. <https://doi.org/10.1111/j.1538-4632.1992.tb00261.x>
- Gilbert M, Pullano G, Pinotti F, Valdano E, Poletto C, Boelle P-Y, D'Ortenzio E, Yazdanpanah Y, Eholie SP, Phram MA, Gutierrez B, Kraemer MUG, Colizza V (2020) Preparedness and vulnerability of African countries against importations of COVID-19: a modelling study. *The Lancet* 395(10227): 14-20 [https://doi.org/10.1016/S0140-6736\(20\)30411-6](https://doi.org/10.1016/S0140-6736(20)30411-6)
- Gong B, Zhang S, Yuan L, Chen KZ (2020) A balance act: minimizing economic loss while controlling novel coronavirus pneumonia. *Journal of Chinese Governance* 1-20. <https://doi.org/10.1080/23812346.2020.1741940>
- Goyal K, Chauhan P, Chhikara K, Gupta P, Singh MP (2020) Fear of COVID 2019: First suicidal case in India! *Asian Journal of Psychiatry* 49:101989 <https://doi.org/10.1016/j.ajp.2020.101989>
- Haque I (2016) Infrastructure development and access to basic amenities in Class-I cities of West Bengal, India: Insights from census data, *Journal of Infrastructure Development*, 8(1):36-84. <https://doi.org/10.1177%2F0974930616640089>

- Haque I, Patel PP (2018) Growth of metro cities in India: trends, patterns and determinants. *Urban Research & Practice* 11(4):338-377. <https://doi.org/10.1080/17535069.2017.1344727>
- Haque I, Rana Md J, Patel PP (2020) Location matters: Unravelling the spatial dimensions of neighbourhood level housing quality in Kolkata, India. *Habitat International* 99:102157. <https://doi.org/10.1016/j.habitatint.2020.102157>
- Harapan H, Itoh N, Yufika A, Winardi W, Keam S, Te H, Megawati D, Hayati Z, Wagner AL, Mudatsir M (2020) Coronavirus disease 2019 (COVID-19): A literature review. *Journal of Infection and Public Health* 13(5):667-673. <https://doi.org/10.1016/j.jiph.2020.03.019>
- Hastie T (2019) gam: Generalized Additive Models. R package version 1.16.1. <https://CRAN.R-project.org/package=gam>
- Ibarra-Vega D (2020) Lockdown, one, two, none, or smart. Modeling containing COVID-19 infection. A conceptual model. *Science of the Total Environment* 730:138917. <https://doi.org/10.1016/j.scitotenv.2020.138917>
- IIPS (International Institute for Population Sciences) and ICF (2017). National Family Health Survey (NFHS-4), 2015-16: India. Mumbai: IIPS
- Jackson M, Huang L, Xie Q, Tiwari RC (2010) A modified version of Moran's I. *International Journal of Health Geographics* 9:33. <https://doi.org/10.1186/1476-072X-9-33>
- Jha R (2020) Post pandemic city planning. Observer Research Foundation, 17th April, 2020. Available at <https://www.orfonline.org/expert-speak/post-pandemic-city-planning-64777/> (accessed on 11th May 2020)
- Jimenez-Sotomayor M, Gomez-Moreno C, Soto-Perez-de-Celis E (2020) Coronavirus, Ageism, and Twitter: An Evaluation of Tweets about Older Adults and COVID-19. *Journal of the American Geriatrics Society* 00:1-5. <https://doi.org/10.1111/jgs.16508>
- Karako K, Song P, Chen Y, Tang W (2020) Analysis of COVID-19 infection spread in Japan based on stochastic transition model. *BioScience Trends* 2020.01482 <https://doi.org/10.5582/bst.2020.01482>
- Kraemer MUG, Yang C, Gutierrez B et al. (2020) The effect of human mobility and control measures on the COVID-19 epidemic in China. *Science* 368:493-497. DOI: 10.1126/science.abb4218
- Kumar S (2020) Predication of Pandemic COVID-19 situation in Maharashtra, India. medRxiv Preprint <https://doi.org/10.1101/2020.04.10.20056697>
- Kundu A (2004) ICT and Human Development: Towards Building a Composite Index for Asia: Realising the Millennium Development Goals, Vi-29. United Nations Development Programme (UNDP). New Delhi: Elsevier.
- Kundu A (1984) Measurement of Urban Processes: A Study in Regionalization. Mumbai: Popular Prakashan.
- Lal N (2020) COVID-19 and India's nowhere people. *The Diplomat*, 1st April, 2020. Available at <https://thediplomat.com/2020/04/covid-19-and-indias-nowhere-people/> (accessed on 11th May 2020)
- Li Q, Guan X, Wu P, et al. (2020) Early transmission dynamics in Wuhan, China, of novel coronavirus-infected pneumonia. *The New England Journal of Medicine* 382:1199-1207. DOI: 10.1056/NEJMoa2001316

- Li B, Yang J, Zhao F, Zhi L, Wang X, Liu L, Bi Z, Zhao Y (2020) Prevalence and impact of cardiovascular metabolic diseases on COVID-19 in China. *Clinical Research in Cardiology* 109: 531-538. <https://doi.org/10.1007/s00392-020-01626-9>
- Lohmueller KE, Pearce CL, Pike M, Lander ES, Hirschhorn JN (2003) Meta-analysis of genetic association studies supports a contribution of common variants to susceptibility to common disease. *Nature Genetics* 33: 177-182 <https://doi.org/10.1038/ng1071>
- Luo W, Majumder MS, Liu D, Poirier C, Mandl KD, Lipsitch M, Santillana M (2020) The role of absolute humidity on transmission rates of the COVID-19 outbreak. *medRxiv Preprint* doi: <https://doi.org/10.1101/2020.02.12.20022467>
- Lupia T, Scabini S, Pinna SM, Perri GD, Rosa FG, Corcoine S (2020) 2019 novel coronavirus (2019- nCoV) outbreak- A new challenge. *Journal of Global Antimicrobial Resistance* 21:22-27. <https://doi.org/10.1016/j.jgar.2020.02.021>
- MacConnachie AA, Collins TC, Seaton RA et al. (2007) Three men, a paint brush and a coronavirus. *International Journal of STD and AIDS* 18:132-133. <https://doi.org/10.1258/095646207779949600>
- Mandal S, Bhatnagar T, Arinaminpathy N, Agarwal A, Chowdhury A, Murhekar M, Gangakhedkar RR, Sarkar S (2020) Prudent public health intervention strategies to control the coronavirus disease 2019 transmission in India: A mathematical model-based approach. *Indian Journal of Medical Research* 151(2): 190-199. DOI: 10.4103/ijmr.IJMR_504_20
- Manderson L, Levine S (2020) COVID-16, Risk, Fear and Fall-out. *Medical Anthropology*. <https://doi.org/10.1080/01459740.2020.1746301>
- McAleer M (2020) Prevention Is Better Than the Cure: Risk Management of COVID-19. *Journal of Risk and Financial Management* 13(3):46. <https://doi.org/10.3390/jrfm13030046>
- McKibbin WJ, Fernando R (2020) The Global Macroeconomic Impacts of COVID-19: Seven Scenarios. CAMA Working Paper No. 19/2020. <https://dx.doi.org/10.2139/ssrn.3547729>
- MoHFW (Ministry of Health and Family Welfare) (2016) District wise health care infrastructure. https://main.mohfw.gov.in/sites/default/files/Final%20RHS%202018-19_0.pdf (accessed 18 April 2020)
- Moran PA (1950) Notes on continuous stochastic phenomena. *Biometrika* 37(1/2):17-23. <https://doi.org/10.1093/biomet/37.1-2.17>
- Mullick J (2020) Covid-19 cases spike in 6 states as migrants return. *Hindustan Times*, 10th June. Available at <https://www.hindustantimes.com/india-news/covid-19-cases-spike-in-6-states-as-migrants-return/story-KdaJIXm3QF3FzOYVxIGBpM.html> (accessed on 14th June 2020)
- Munoz MD, Meyer T (2020) Infectious Diseases and Football- Lessons not only from COVID-19. *Science and Medicine in Football* 4(2):85-86. <https://doi.org/10.1080/24733938.2020.1749422>
- Nacoti M, MD, Ciocca A et al. (2020) At the Epicenter of the COVID-19 Pandemic and Humanitarian Crises in Italy: Changing Perspectives on Preparation and Mitigation. *Innovations in Care Delivery*. 10.1056/CAT.20.0080

- NDMA (National Disaster Management Authority) (2020) Hotspot' (Red Zone) classification'. Available at <https://ndma.gov.in/images/covid/MoHFW-Letter-States-reg-containment-of-Hotspots.pdf> (accessed on 11th May 2020)
- Nicola M, Alsafi Z, Sohrabi C et al. (2020) The Socio-Economic Implications of the Coronavirus and COVID-19 Pandemic: A Review. *International Journal of Surgery*. <https://doi.org/10.1016/j.ijssu.2020.04.018>
- Nidheesh MK (2020) Kerala reopens without communal transmission despite covid rise. *livemint*, 9th June. Available at <https://www.livemint.com/news/india/kerala-reopens-without-communal-transmission-despite-covid-rise-11591647068759.html> (accessed on 14th June 2020)
- Office of the Registrar General of India (2011) Census of India 2011. Ministry of Home Affairs, Government of India. New Delhi, India.
- Oliver N, Lepri B, Sterly H et al. (2020) Mobile phone data for informing public health actions across the COVID-19 pandemic lifecycle. *Science Advances*. 10.1126/sciadv.abc0764
- Ord, J.K. and Getis, A. (1995), Local Spatial Autocorrelation Statistics: Distributional Issues and an Application. *Geographical Analysis*, 27:286-306. doi:10.1111/j.1538-4632.1995.tb00912.x
- Paital B, Das K, Parida SK (2020) Inter nation social lockdown versus medical care against COVID-19, a mild environmental insight with special reference to India. *Science of the Total Environment* 728:138914. <https://doi.org/10.1016/j.scitotenv.2020.138914>
- Pandey G, Chaudhury P, Gupta R, Pal S (2020) SEIR and Regression Model based COVID-19 outbreak predictions in India. *Populations and Evolution* (preprint) arXiv:2004.00958 [q-bio.PE]
- Parnell D, Widdop P, Bond A, et al. (2020) COVID-19, networks and sport. *Managing Sport and Leisure* 1-7. <https://doi.org/10.1080/23750472.2020.1750100>
- Prata DN, Rodrigues W, Bermejo PH (2020) Temperature significantly changes COVID-19 transmission in (sub) tropical cities of Brazil. *Science of the Total Environment* 729:138862. <https://doi.org/10.1016/j.scitotenv.2020.138862>
- Prem K, Liu Y, Russell TW, Kucharski AJ, Eggo RM, Davies N et al. (2020) The effect of control strategies to reduce social mixing on outcomes of the COVID-19 epidemic in Wuhan, China: a modelling study. *The Lancet* 5:e261-270. [https://doi.org/10.1016/S2468-2667\(20\)30073-6](https://doi.org/10.1016/S2468-2667(20)30073-6)
- Qi H, Xiao S, Shi R, et al. (2020) COVID-19 transmission in Mainland China is associated with temperature and humidity: A time-series analysis. *Science of the Total Environment* 728:138778. <https://doi.org/10.1016/j.scitotenv.2020.138778>
- Raffaeta R (2020) Another Day in Dystopia. Italy in the Time of COVID-19. *Medical Anthropology* 1-3. <https://doi.org/10.1080/01459740.2020.1746300>
- Regmi MB (2020) COVID-19 prompts rethinking of mobility and city planning. United Nations ESCAP, 17th April, 2020. Available at <https://www.unescap.org/blog/covid-19-prompts-rethinking-mobility-and-city-planning> (accessed on 11th May 2020)
- Requia WJ, Kondo EK, Adams MD et al. (2020) Risk of the Brazilian health care system over 5572 municipalities to exceed health care capacity due to the 2019 novel coronavirus (COVID-19). *Science of the Total Environment* 730:139144. <https://doi.org/10.1016/j.scitotenv.2020.139144>

- Rocklöv J, Sjödin H (2020) High population densities catalyze the spread of COVID-19. *Journal of Travel Medicine* 1-2:taaa03. doi: 10.1093/jtm/taaa038
- Rothan HA, Byrareddy SN (2020) The epidemiology and pathogenesis of coronavirus disease (COVID-19) outbreak. *Journal of Autoimmunity* 109: 102433 <https://doi.org/10.1016/j.jaut.2020.102433>
- Rowan NJ, Laffey JG (2020) Challenges and solutions for addressing critical shortage of supply chain for personal and protective equipment (PPE) arising from Coronavirus disease (COVID-19) pandemic- case study from the Republic of Ireland. *Science of the Total Environment* 725:138532. <https://doi.org/10.1016/j.scitotenv.2020.138532>
- Sajadi MM, Habibzadeh P, Vintzileos A, Shokouhi S, Miralles-Wilhelm F, Amoroso A (2020) Temperature and latitude analysis to predict potential spread and seasonality for COVID-19. SSRN. <https://dx.doi.org/10.2139/ssrn.3550308>
- Sardar T, NadimSk S, Chattopadhyay J (2020) Assessment of 21 Days Lockdown Effect in Some States and Overall India: A Predictive Mathematical Study on COVID-19 Outbreak. *Populations and Evolution* (preprint) <https://arxiv.org/abs/2004.03487> [q-bio.PE]
- Sharfuddin S (2020) A CHOGM under COVID cloud. *The Round Table, The Commonwealth Journal of International Affairs*, 1-2. <https://doi.org/10.1080/00358533.2020.1741892>
- Shi Y, Yu X, Zhao H, Wang H, Zhao R, Sheng J (2020) Host susceptibility to severe COVID-19 and establishment of a host risk score: findings of 487 cases outside Wuhan. *Critical Care* 24:108 <https://doi.org/10.1186/s13054-020-2833-7>
- Simha A, Prasad RV, Narayana S (2020) A simple Stochastic SIR model for COVID 19 Infection Dynamics for Karnataka: Learning from Europe. *Populations and Evolution* (preprint) <https://arxiv.org/abs/2003.11920> [q-bio.PE]
- Singh AK, Singh A, Shaikh A, Singh R, Misra (2020) Chloroquine and hydroxychloroquine in the treatment of COVID-19 with or without diabetes: A systematic search and a narrative review with a special reference to India and other developing countries. *Diabetes & Metabolic Syndrome: Clinical Research & Reviews* 14(3):241-246 <https://doi.org/10.1016/j.dsx.2020.03.011>
- Singh R, Adhikari R (2020) Age-structured impact of social distancing on the COVID-19 epidemic in India. Available at <https://github.com/rajeshrinet/pyross> (accessed on 11th May 2020)
- Singh S (2020) Explained: Indian migrants, across India. *The Indian Express*, 29th April, 2020. Available at <https://indianexpress.com/article/explained/coronavirus-india-lockdown-migran-workers-mass-exodus-6348834/> (accessed 11th May 2020)
- Sominsky L, Walker DW, Spencer SJ (2020) One size does not fit all- Patterns of vulnerability and resilience in the COVID-19 pandemic and why heterogeneity of disease matters. *Brain, Behaviour and Immunity* <https://doi.org/10.1016/j.bbi.2020.03.016>
- Soreide K, Hallet J, Matthews JB, et al. (2020) Immediate and long-term impact of the COVID-19 pandemic on delivery of surgical services. *BJS Society*. <https://doi.org/10.1002/bjs.11670>

- Srinivasa Rao ASR, Vazquez JA (2020) Identification of COVID-19 can be quicker through artificial intelligence framework using a mobile phone-based survey when cities and towns are under quarantine. *Infection Control and Hospital Epidemiology*. <https://doi.org/10.1017/ice.2020.61>
- The Lancet (2020) Indian under COVID-19 lockdown. *The Lancet* 395. [https://doi.org/10.1016/S0140-6736\(20\)30938-7](https://doi.org/10.1016/S0140-6736(20)30938-7)
- Ting DSW, Carin L, Dzau V, Wong TY (2020) Digital technology and COVID-19. *Nature Medicine* 26: 459-461. <https://doi.org/10.1038/s41591-020-0824-5>
- Tobias A (2020) Evaluation of the lockdowns of the SARS-CoV-2 epidemic in Italy and Spain after one month follow up. *Science of the Total Environment* 725:138539. <https://doi.org/10.1016/j.scitotenv.2020.138539>
- Tomar A, Gupta N (2020) Prediction for the spread of COVID-19 in India and effectiveness of preventive measures. *Science of the Total Environment* 728:138762. <https://doi.org/10.1016/j.scitotenv.2020.138762>
- Torales J, O'Higgins M, Castaldelli-Maia JM et al. (2020) The outbreak of COVID-19 coronavirus and its impact on global mental health. *International Journal of Social Psychiatry* 66(4):1-4. <https://doi.org/10.1177%2F0020764020915212>
- Van den Berg R (2020) How will COVID-19 affect urban planning? The CityFix, 10th April, 2020. Available at <https://thecityfix.com/blog/will-covid-19-affect-urban-planning-roger-van-den-berg/> (accessed on 11th May 2020)
- Wang J, Luo Q, Chen R, Chen T, Li J (2020) Susceptibility Analysis of COVID-19 in Smokers Based on ACE2. Preprints 2020030078 doi: 10.20944/preprints202003.0078.v1
- Wang L, Wang Y, Ye D, Liu Q (2020) A review of the 2019 novel Coronavirus (SARS-CoV-2) based on current evidence. *International Journal of Antimicrobial Agents* 105948. <https://doi.org/10.1016/j.ijantimicag.2020.105948>
- Wang Y, Di Y, Ye J et al. (2020) Study on the public psychological states and its related factors during the outbreak of coronavirus disease 2019 (COVID-19) in some regions of China. *Psychology, Health and Medicine*. <https://doi.org/10.1080/13548506.2020.1746817>
- Wang Z, Tang K (2020) Combating COIV-19: Health equity matters. *Nature Medicine* 26:458-464. <https://doi.org/10.1038/s41591-020-0823-6>
- Wang J, Tang K, Feng K, Lv W (2020) High Temperature and High Humidity Reduce the Transmission of COVID-19. SSRN. Populations and Evolution. arXiv:2003.05003 [q-bio.PE]
- Wei T, Simko V (2017). R package "corrplot": Visualization of a Correlation Matrix (Version 0.84). Available from <https://github.com/taiyun/corrplot>
- WHO (World Health Organisation) (2020a) (World Health Organisation) (2020) Newsroom/Coronavirus (Covid-19). <https://www.who.int/news-room/q-a-detail/q-a-coronaviruses> (accessed on 8th April 2020)
- WHO (World Health Organisation) (2020b) Coronavirus disease 2019 (COVID-19) Situation Report-73. World Health Organisation, Geneva. Available at <https://www.who.int/docs/default-source/coronaviruse/situation-reports/20200402-sitrep-73-covid-19.pdf> (accessed on 8th April 2020)

- Willian J, King AJ, Djebbari F, Turner GDH, Royston DJ, Pavord S, Collins GP, Peniket A (2020) Assessing the Impact of Lockdown: Fresh Challenges for the Care of Haematology Patients in the COVID-19 Pandemic. *British Journal of Haematology* <https://doi.org/10.1111/bjh.16782>
- Wu Y, Jing W, Liu J, et al. (2020) Effects of temperate and humidity on the daily new cases and new deaths of COVID-19 in 166 countries. *Science of the Total Environment* 729:139051. <https://doi.org/10.1016/j.scitotenv.2020.139051>
- Wu X, Nethery RC, Sabath MB, Braun D, Dominici F (2020) Exposure to air pollution and COVID-19 mortality in the United States. *medRxiv* (preprint) <https://doi.org/10.1101/2020.04.05.20054502>
- Xie J, Zhu Y (2020) Association between ambient temperature and COVID-19 infection in 122 cities from China. *Science of the Total Environment* 724:138201. <https://doi.org/10.1016/j.scitotenv.2020.138201>
- Yang J, Zheng Y, Gou X, Pu K, Chen Z, Guo Q, Ji R, Wang H, Wang Y, Zhou Y (2020) Prevalence of comorbidities and its effects in patients infected with SARS-CoV-2: a systematic review and meta-analysis. *International Journal of Infectious Diseases* 94: 91-95 <https://doi.org/10.1016/j.ijid.2020.03.017>
- Zhang J, Litvinova M, Liang Y, Wang Y, Wang W, Zhao S, Wu Q, Merler S, Viboud C, Vespignani A, Ajelli M, Yu H (2020) Changes in contact patterns shape the dynamics of the COVID-19 outbreak in China. *Science* 1-12. [10.1126/science.abb8001](https://doi.org/10.1126/science.abb8001)
- Zhao J, Yang Y, Huang H, Li D, Gu D, Lu X, Zhang Z, Liu L, Liu T, Liu Y, He Y, Sun B, Wei M, Yang G, Wang X, Zhang L, Zhou X, Xing M, Wang PG (2020) Relationship between the ABO Blood Group and the COVID-19 Susceptibility. *medRxiv Preprint* <https://doi.org/10.1101/2020.03.11.20031096>
- Zhong, B. L., Luo, W., Li, H. M., Zhang, Q. Q., Liu, X. G., Li, W. T., & Li, Y. (2020). Knowledge, attitudes, and practices towards COVID-19 among Chinese residents during the rapid rise period of the COVID-19 outbreak: a quick online cross-sectional survey. *International Journal of Biological Sciences* 16(10):1745. doi:10.7150/ijbs.45221
- Zhou C, Su F, Pei T, et al. (2020) COVID-19: Challenges to GIS with Big Data. *Geography and Sustainability* 1(1):77-87. <https://doi.org/10.1016/j.geosus.2020.03.005>

Figure 1: India, with different States, Union Territories and their districts (Note: Administrative divisions are shown as per the Indian Population Census of 2011)

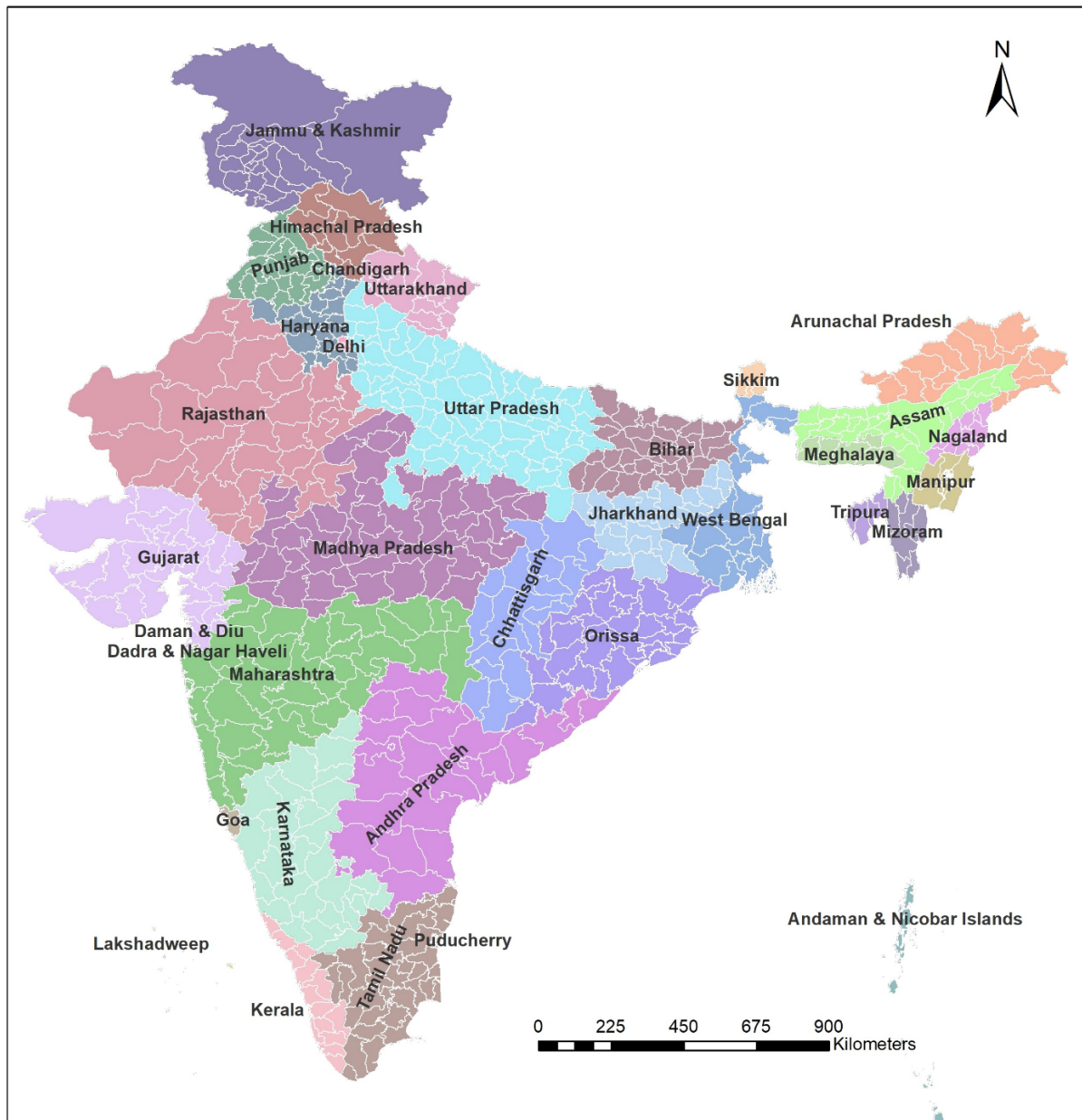


Figure 2: The outline devised for the analysis

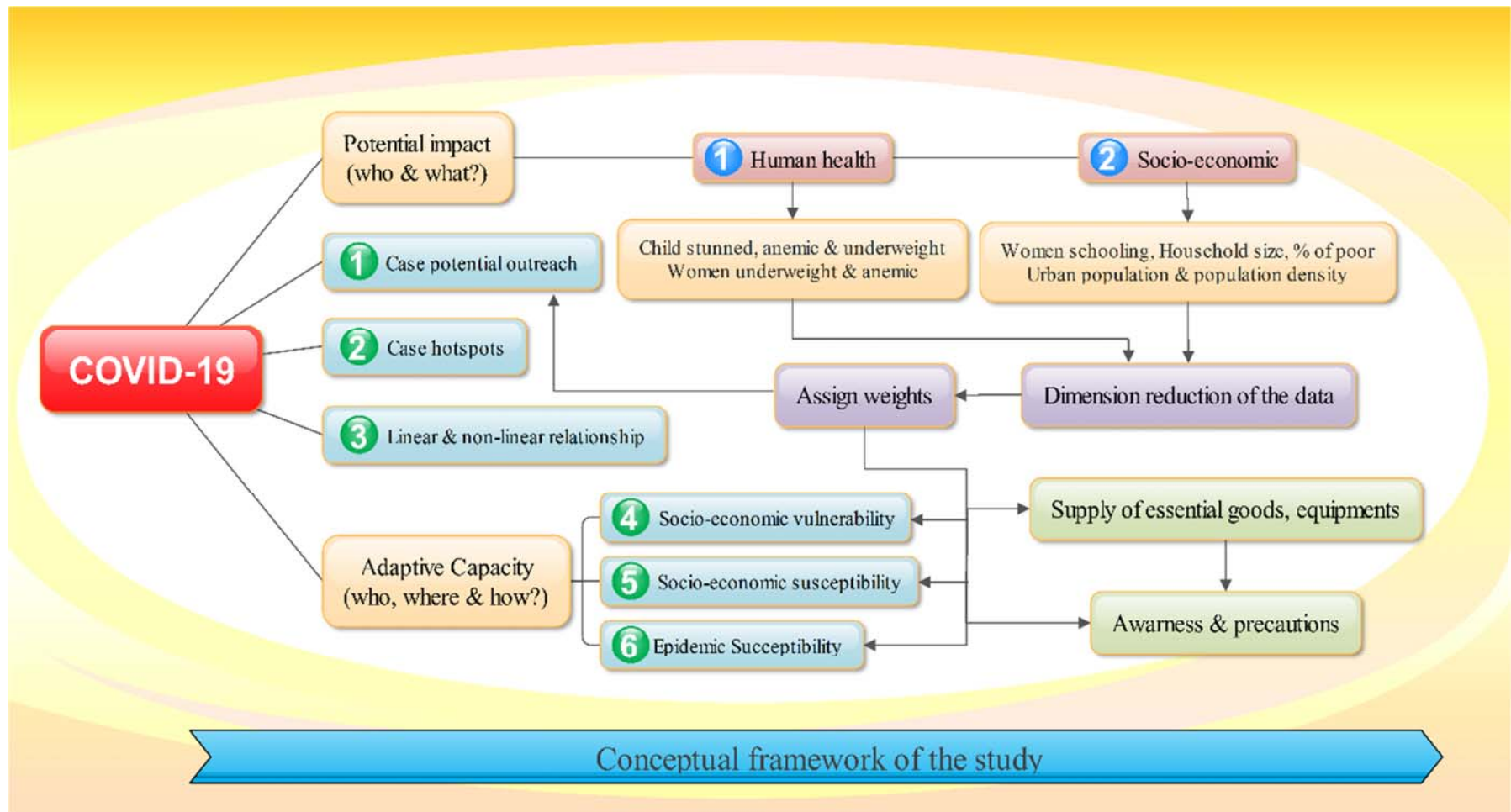


Figure 3: Spread of COVID-19 across India's districts, with the situation as on: (A) On 23rd March, 2020; (B) 29th March, 2020; (C) 10th April, 2020, and (D) 18th April, 2020

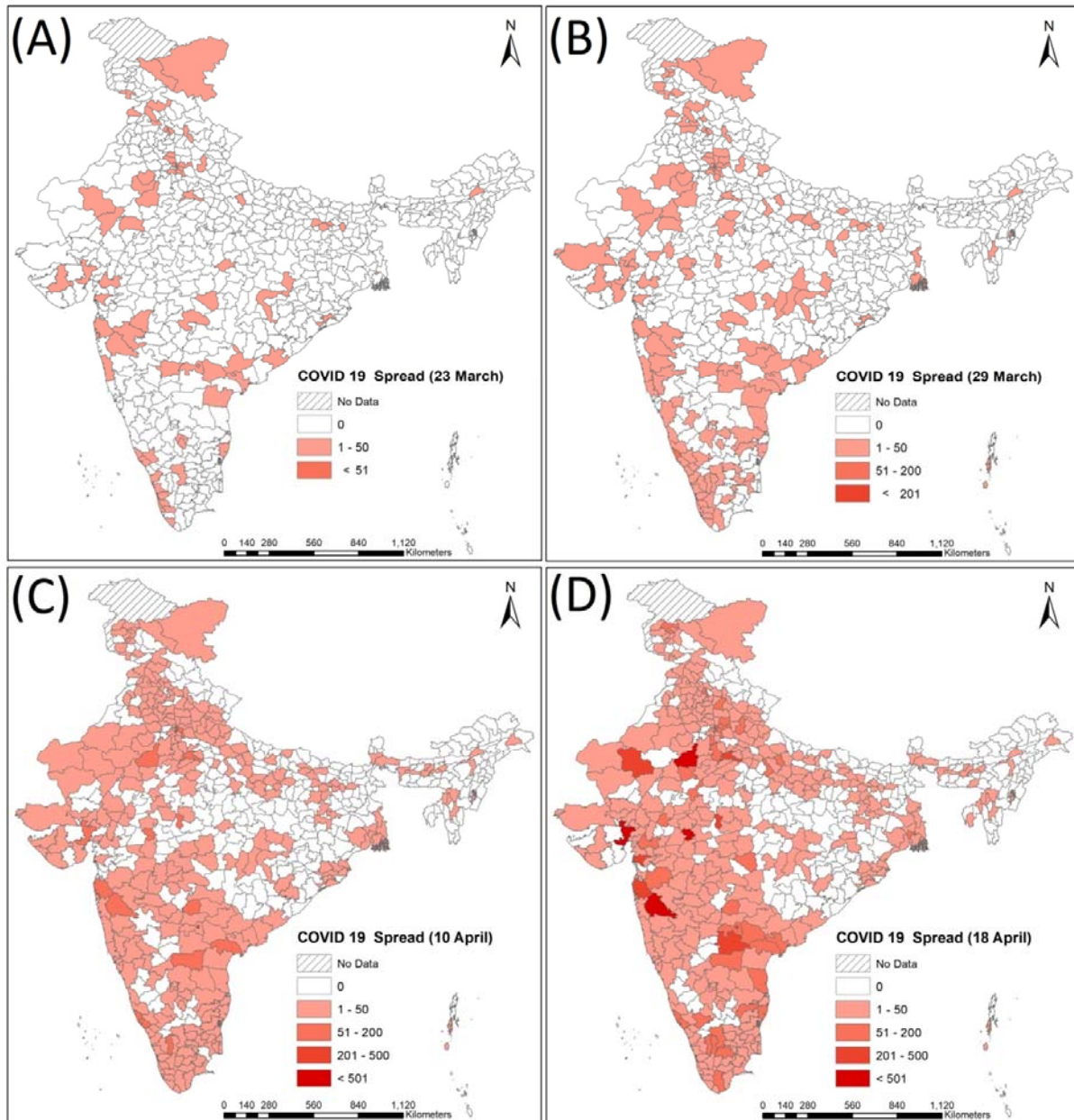


Figure 4: Relation between incidence of COVID-19 cases and distance from nearest international airport

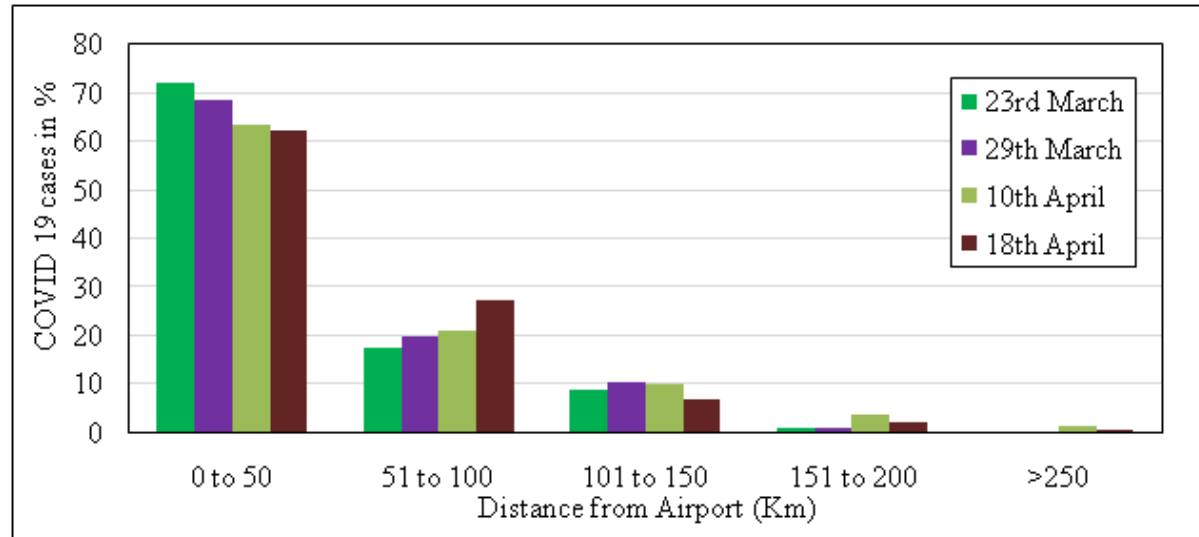


Figure 5: Airport passenger volume link with COVID-19 incidence across India

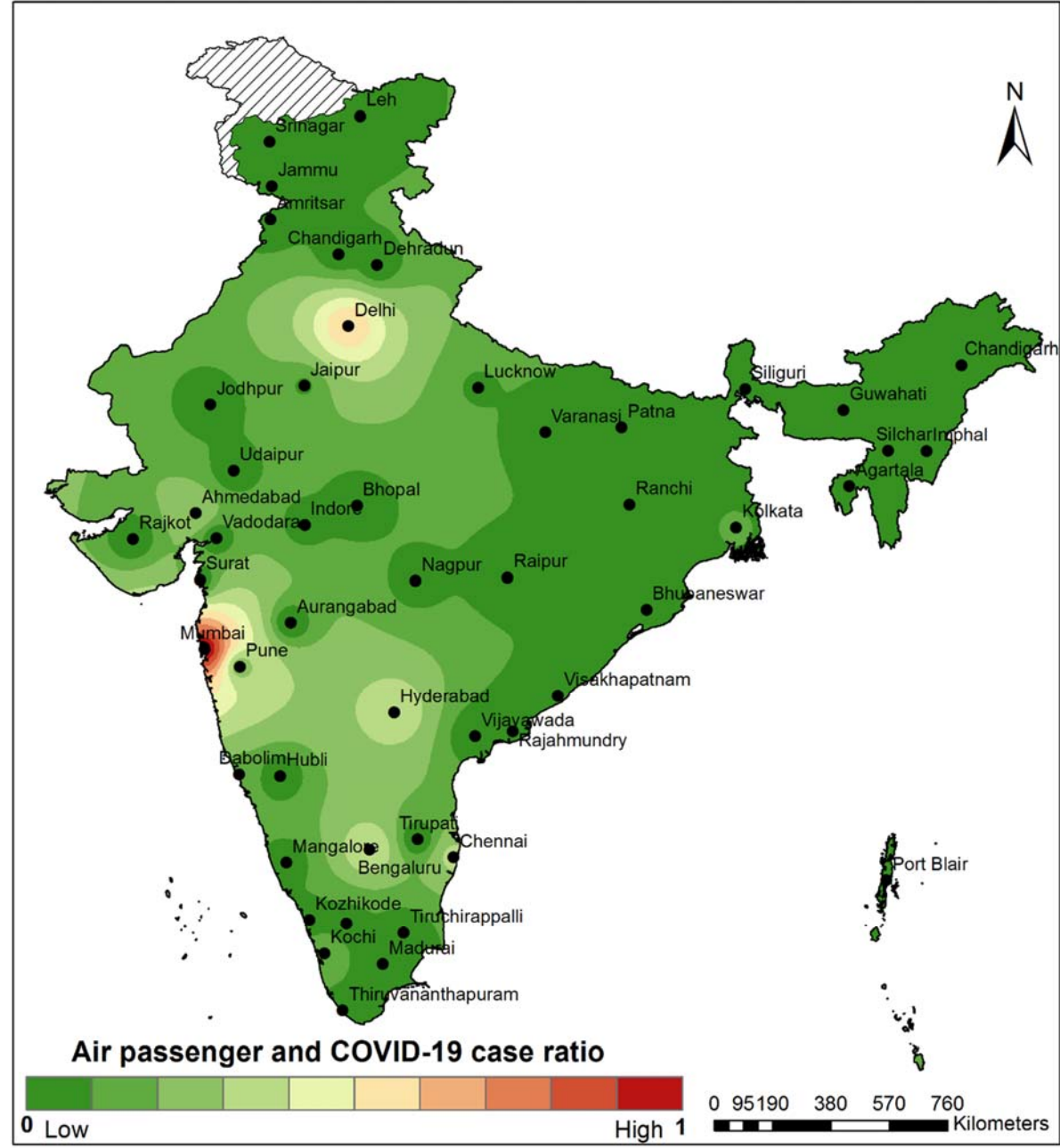


Figure 6: Hotspot and Cold spot clusters of COVID-19 cases across India's districts, with the situation as on: (A) 23rd March, 2020; (B) 29th March, 2020; (C) 10th April, 2020, and (D) 18th April, 2020

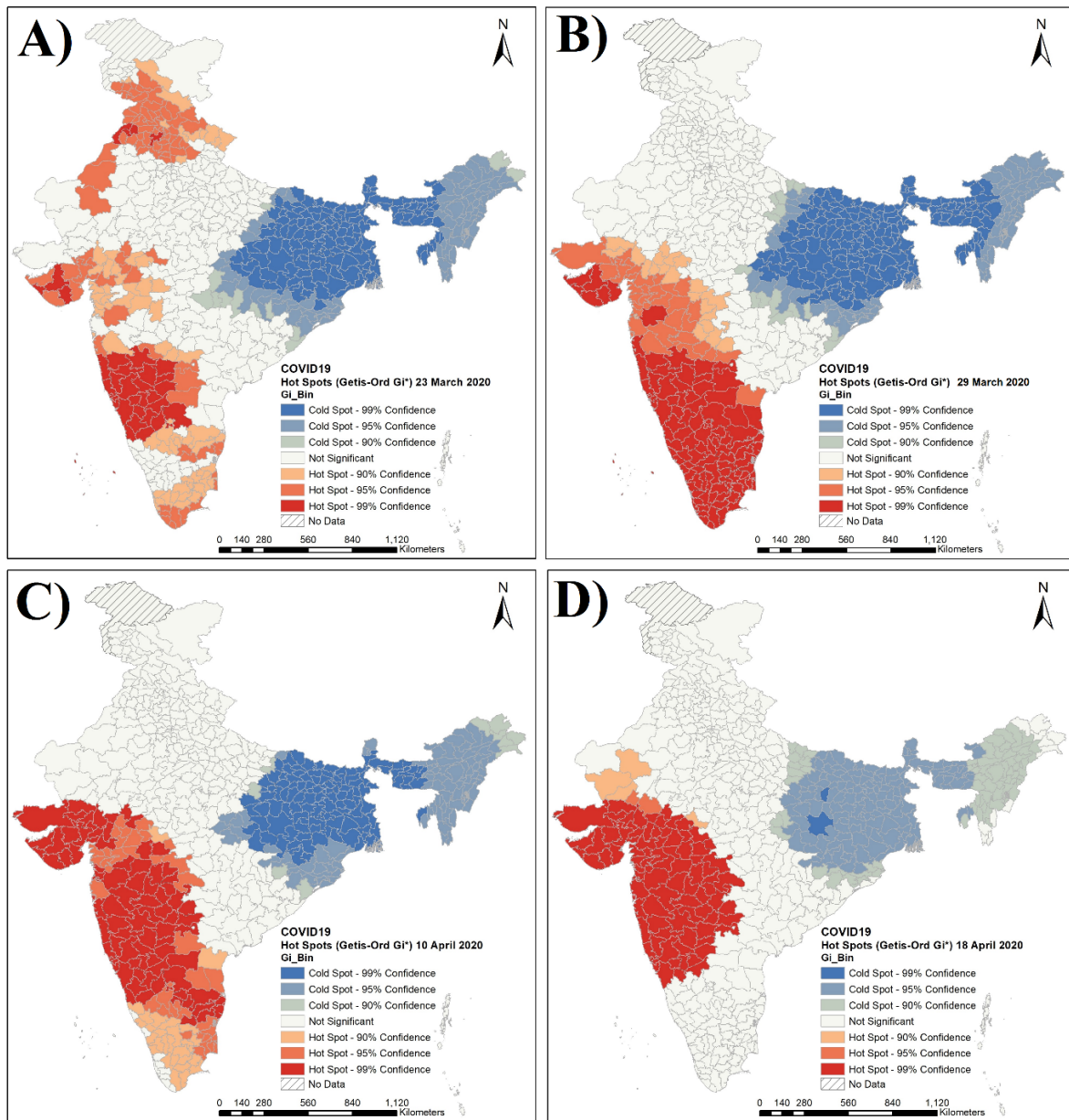


Figure 7: Moran's I value-based clustering of COVID-19 cases across India's districts, with the situation as on: (A) 23rd March, 2020; (B) 29th March, 2020; (C) 10th April, 2020, and (D) 18th April, 2020

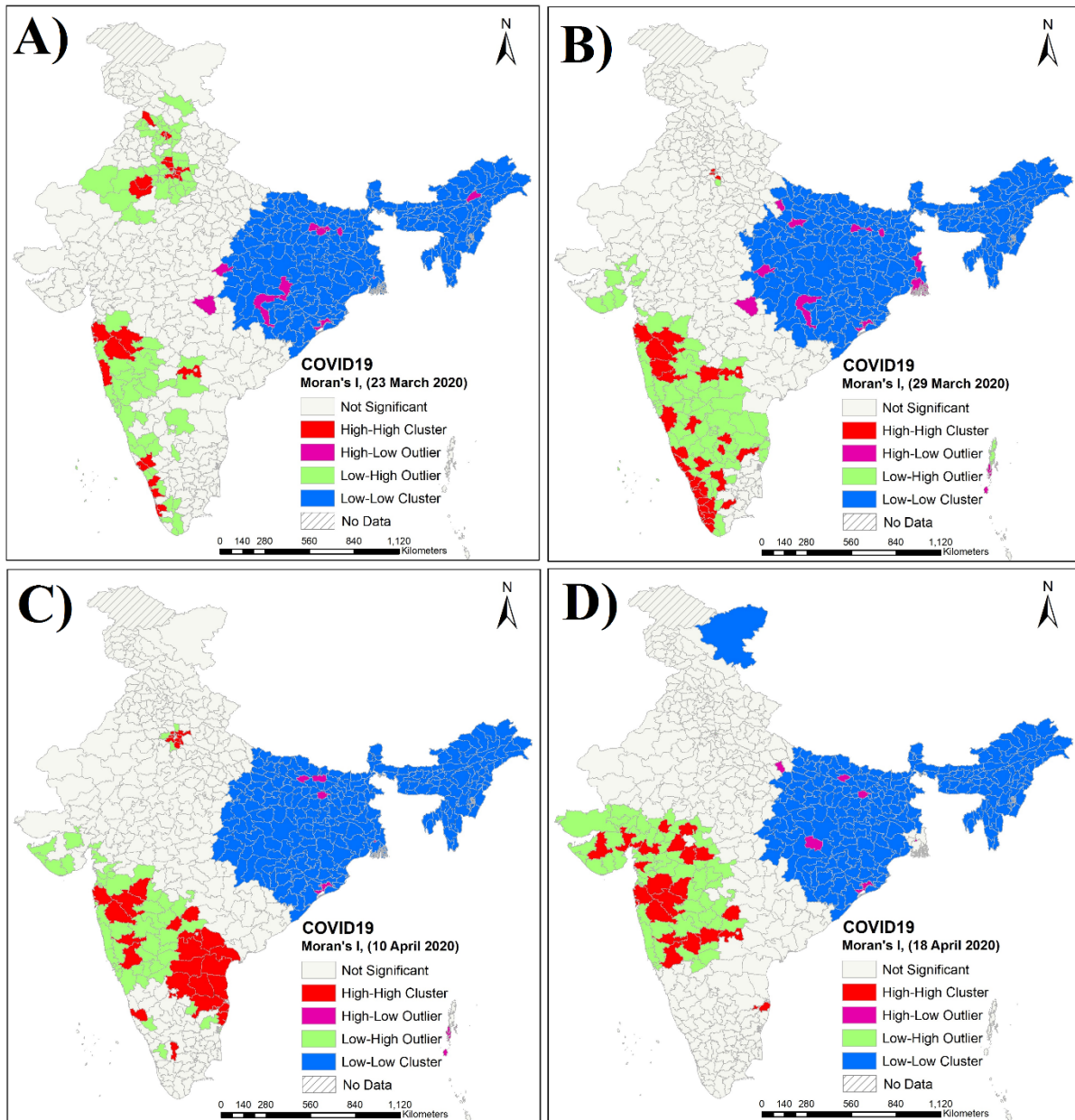


Figure 8: Potential outreach zones of the COVID-19 virus across India's districts, with the situation as on: (A) 23rd March, 2020; (B) 29th March, 2020; (C) 10th April, 2020, and (D) 18th April, 2020

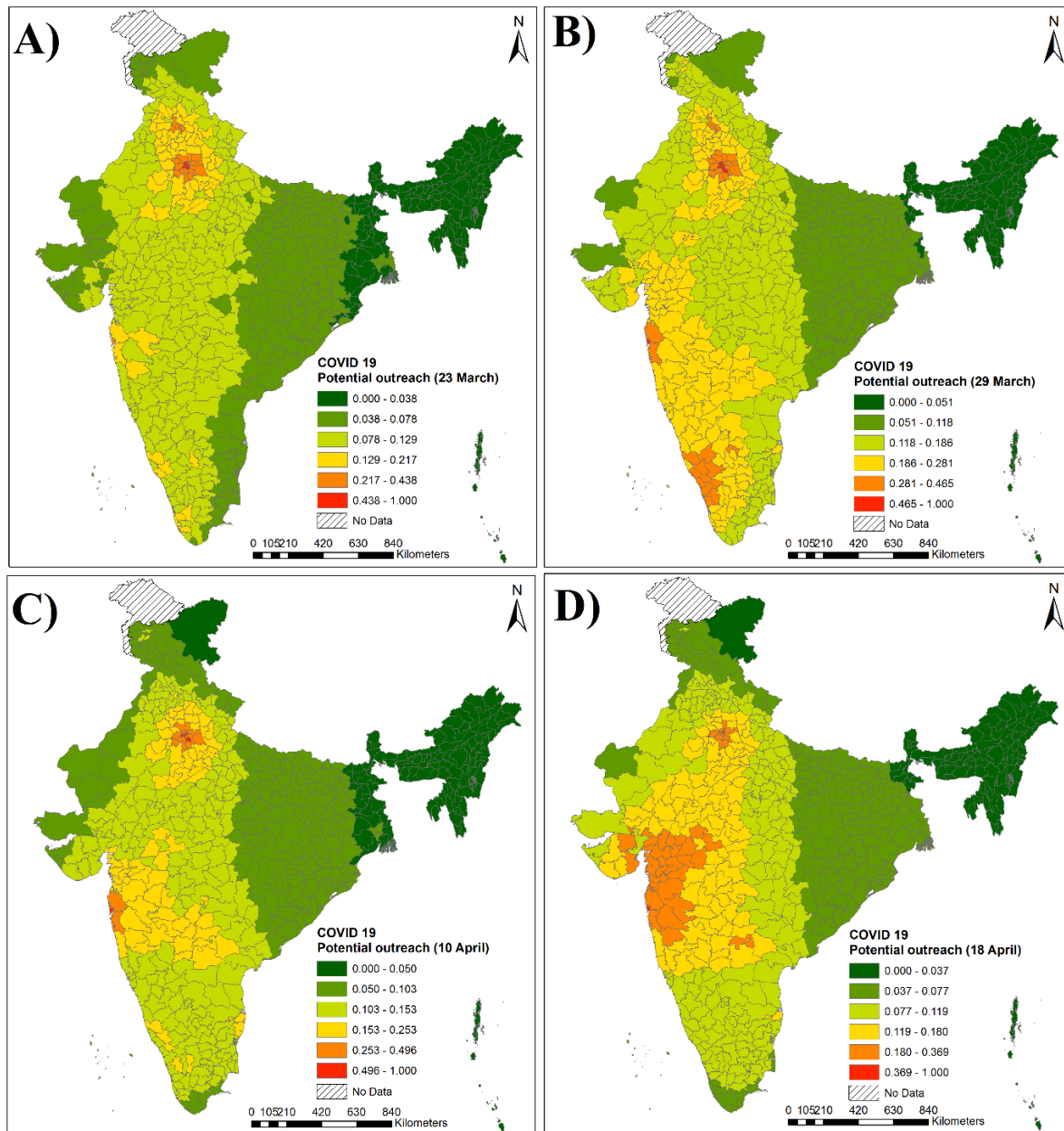


Figure 9: District-level epidemiological susceptibility to COVID-19 in India

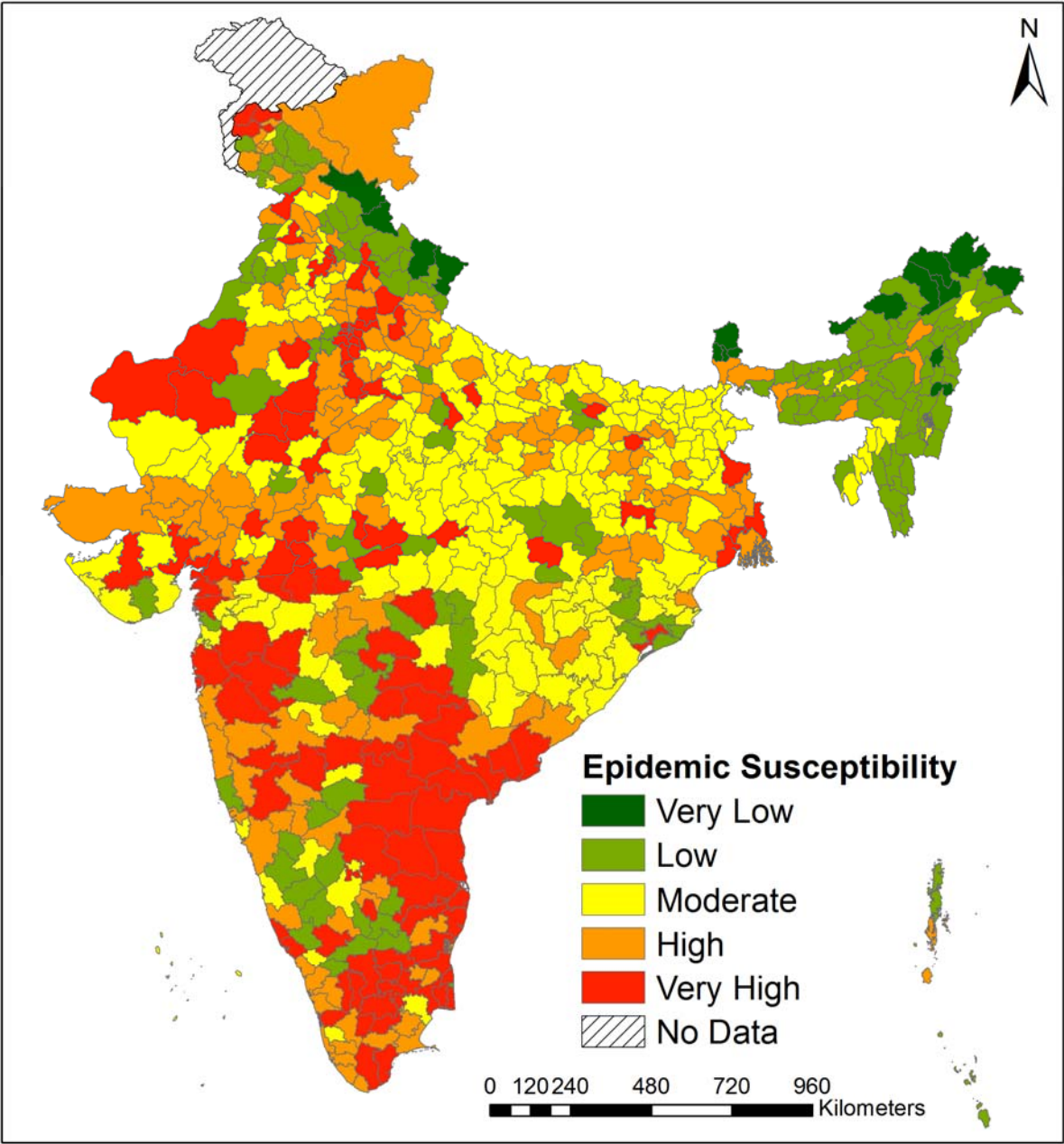


Figure 10: District-level socioeconomic susceptibility to COVID-19 in India

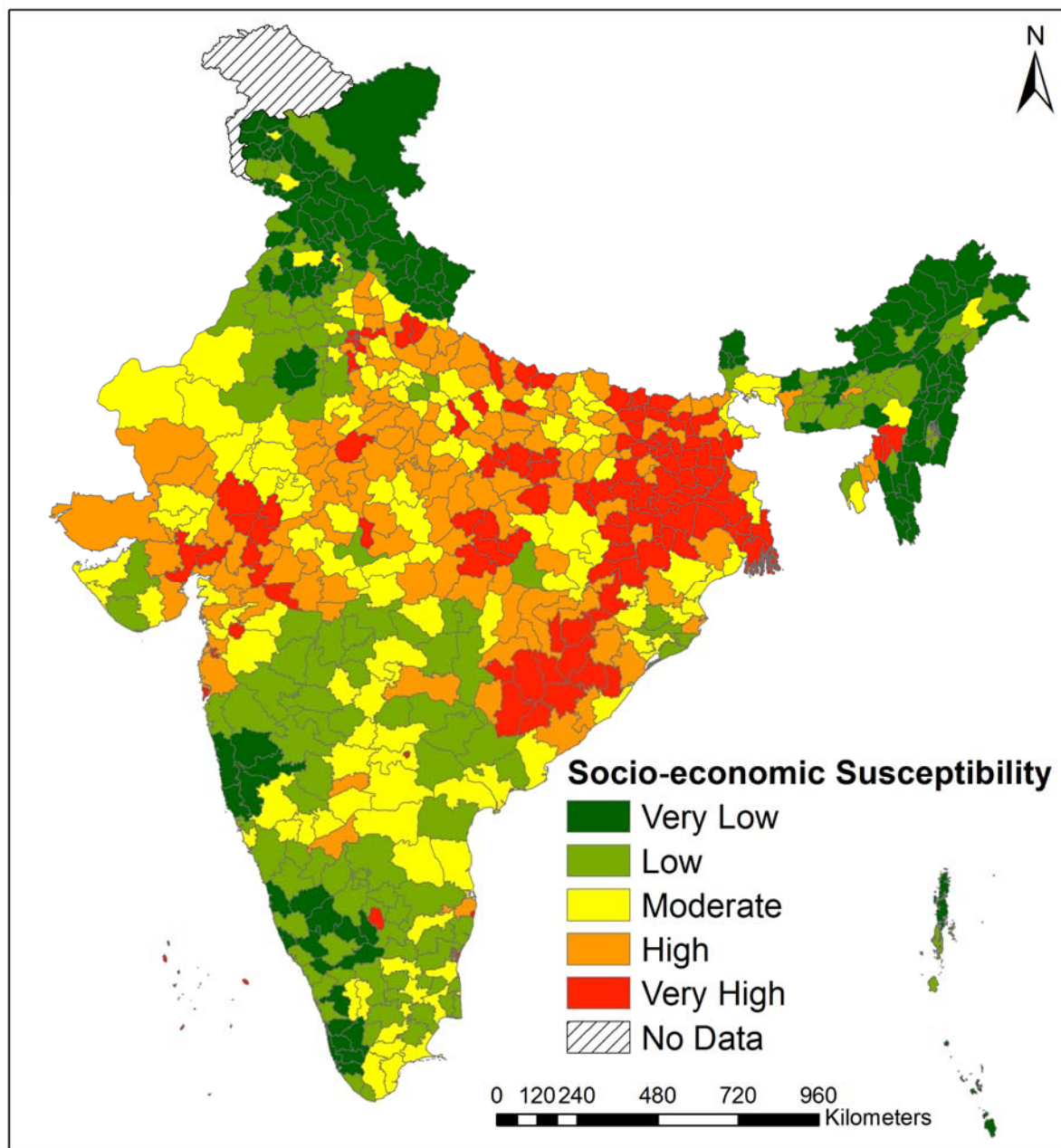


Figure 11: District-level socioeconomic vulnerability to COVID-19 in India

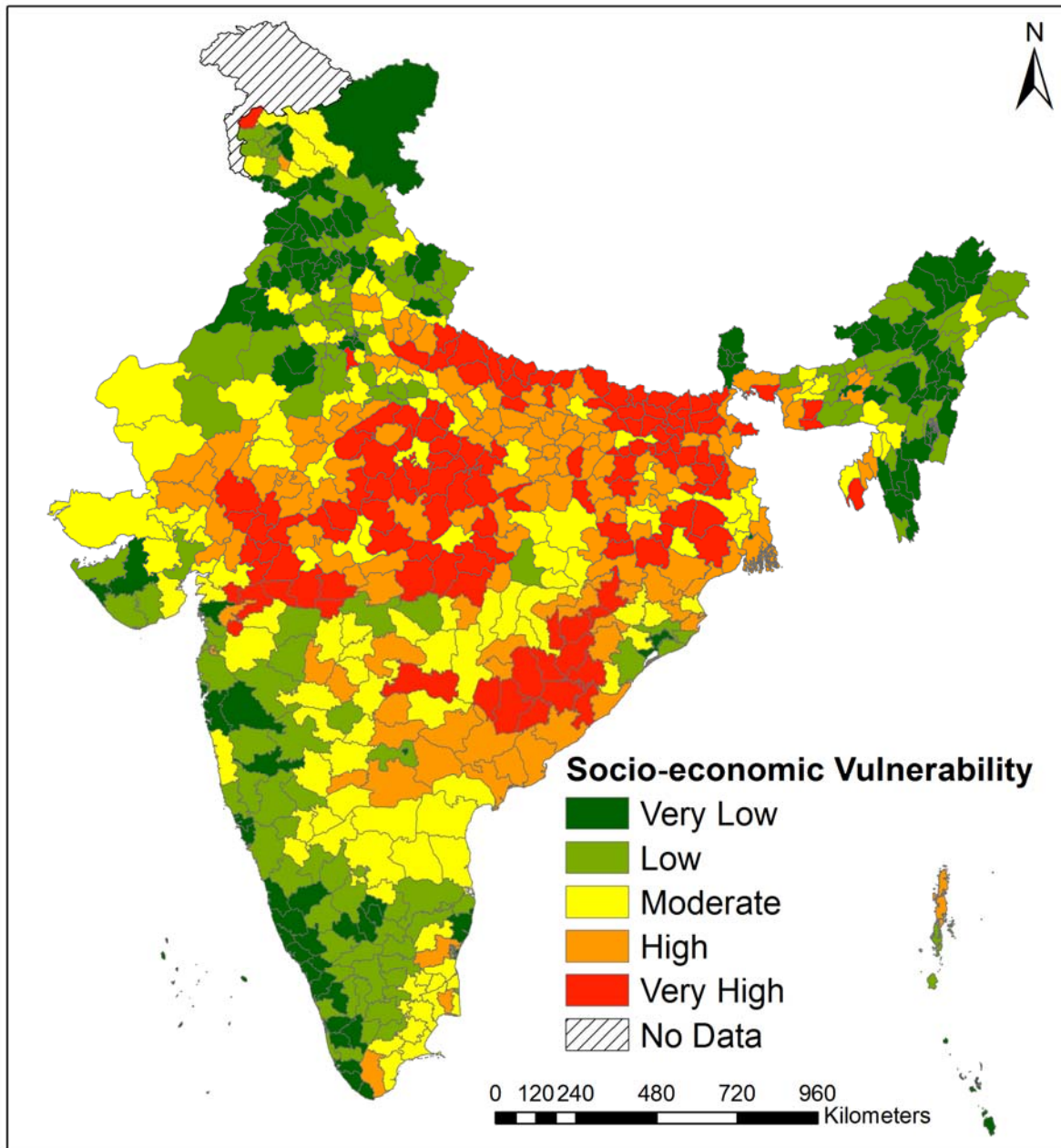


Figure 12: Enumerated relations between the district-level incidence of COVID-19 cases and its climatic, demographic and socioeconomic correlates

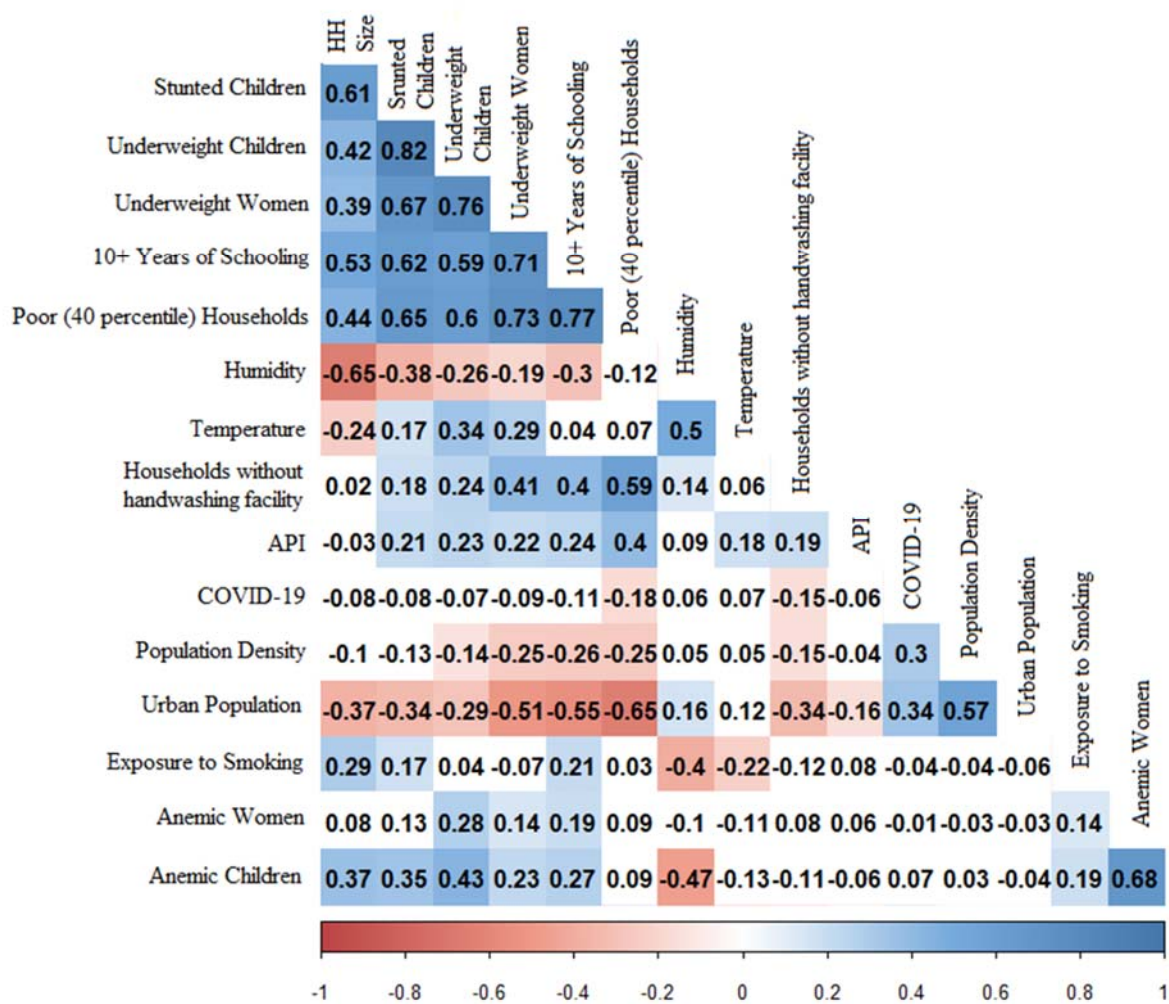


Figure 13: Results of the Generalized Additive Model

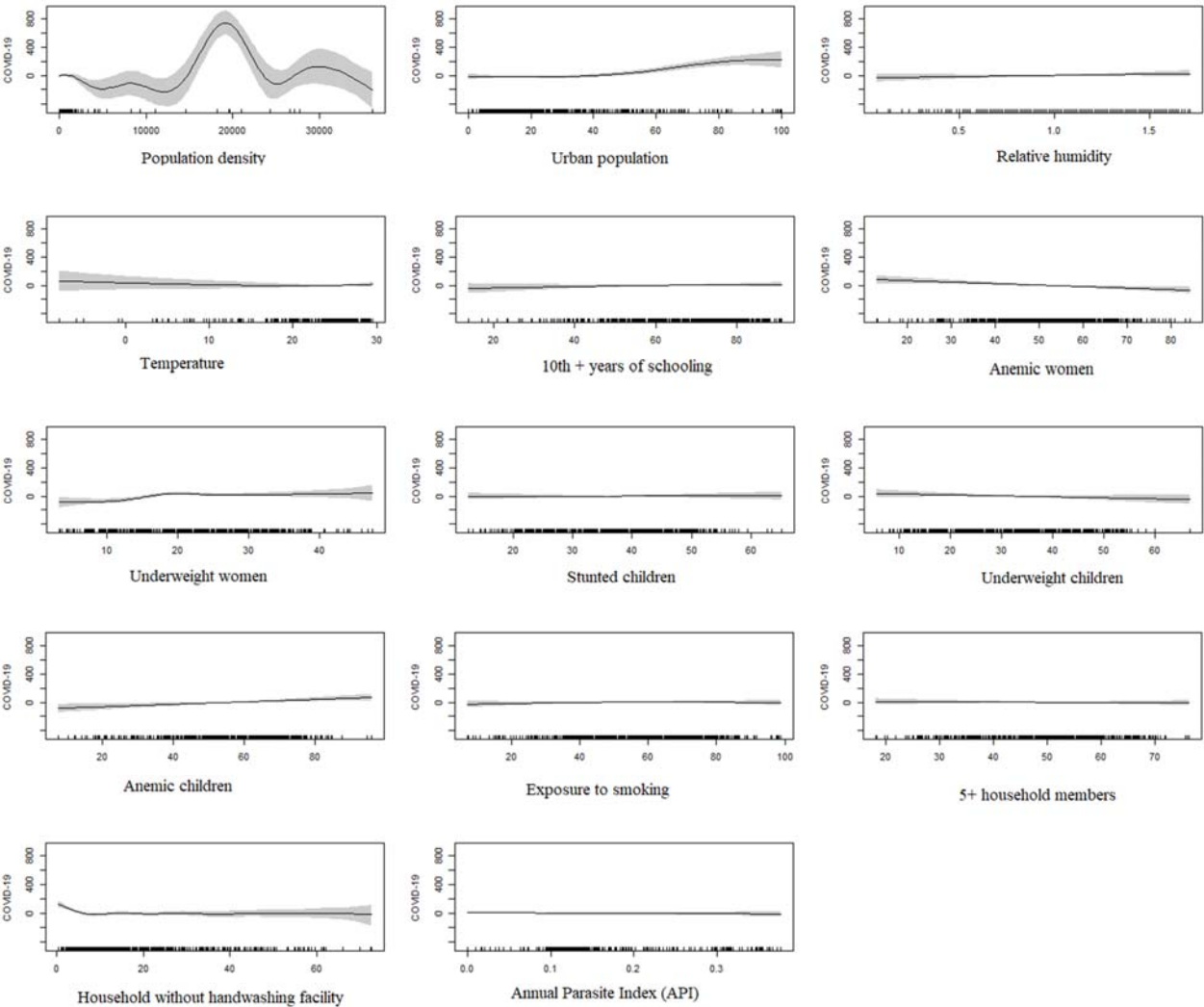


Table 1: Indicators for estimating the susceptibility of infection by COVID-19

Indicators	Assumptions/References	Data sources	Weight (%) - Epidemiological susceptibility	Weight (%) - Socioeconomic susceptibility
Population density (per sq. km)	Densely populated areas raise the risk of greater physical contact among people. As the SARS-COV-2 is a highly contagious virus, population density is an unavoidable measure (Rocklöv and Sjödin 2020)	Office of the Registrar General of India (2011)	1	7
Percentage of the total population living in urban areas	Urban areas are particularly most vulnerable as they are the initial centres of the outbreak. The prevalence of slums within cities especially raises the risk of this viral transmission (Coudhry and Avindandan 2020).	Office of the Registrar General of India (2011)	1	7
Percentage of women having 10th standard education and above	Proxy parameter for the awareness about the COVID-19 spreading. It is assumed that if an area has a higher proportion of women 10 or more years of schooling, then overall household awareness will be higher and this can lower transmission (Zhong et al. 2020).	IIPS and ICF (2017)	1	7
Percentage of households exposed to smoking	Smoking or being exposed to smoking in any form can elevate the risk of infection from COVID-19 (WHO 2020).	IIPS and ICF (2017)	1	7
Percentage of anaemic women	Undernourishment or micronutrient deficiency in the body raises the risk of infection from any communicable diseases (França et al. 2009).	IIPS and ICF (2017)	1	7
Percentage of underweight women		IIPS and ICF (2017)	1	7
Percentage of children stunting		IIPS and ICF (2017)	1	7
Percentage of children underweight		IIPS and ICF (2017)	1	7
Percentage of children anaemia		IIPS and ICF (2017)	1	7
Average temperature (in °C)	Both lower temperature and humidity increases the risk of spreading COVID-19 (Sajadi et al. 2020; Wang et al. 2020)	WorldClim dataset	1	7
Relative humidity (%)		NASA GES	1	7
Number of reported cases	A higher number of cases already reported from an area may push it towards becoming a hotspot for further spreading	MoHFW, GoI, 2020	87	0
Percentage of large household size (5+)	Congestions of household members may elevate the risk of infections.	IIPS and ICF (2017)	1	7
Percentage of households with places for hand washing	Hand-washing would reduce the risk infection.	IIPS and ICF (2017)	1	7
Malarious districts/ chloroquine drug distributed areas	Generally, malaria cases are calculated by Annual Parasite Index (API) [i.e. Plasmodium vivax (Pv) and Plasmodium falciparum (Pf)]. In malarious districts, chloroquine is the main drug used to alleviate symptoms. Since recent research proves that chloroquine has the potentiality to treat patients infected by SARS-CoV-2/COVID-19 (Colson et al. 2020; Gao et al. 2020), it is assumed that immunity capabilities of people in malarious districts shall be higher those in other areas	NVBDCP and Battle et al. 2015	0	9

Note: Zero (0), if specified in the columns of weights, indicates the exclusion of that variable from the model

Source: Prepared by the authors

Table 2: Indicators for estimating the vulnerability to infection by COVID-19

Indicators	Assumptions/References	Data sources	Weight (%)
Percentage of children stunted	Immunity is one of the major drivers to counter the COVID-19 virus. As no district-level indicators of immunity are available, we have relied on other proxy health indicators (e.g. França et al. 2009).	IIPS and ICF (2017)	8
Percentage of children underweight			8
Percentage of anaemic children			8
Percentage of women underweight			8
Percentage of anaemic women			8
No. of diabetic patients and those having heart ailments			9
Percentage of women with 10th standard education or lower	Educated people are not only aware about how to prevent infection but also can deal with asymptomatic and mild symptomatic persons better than uneducated counterparts. They can also help others (family members and neighbours) to deal with an infected person more effectively. Furthermore, educated people are likely to have better attitudes and practices for enabling COVID-19 prevention (Zhong et al. 2020).	IIPS and ICF (2017)	8
Percentage of aged population (60+ years)	Aged population have a higher risk of death after infection (Dowd et al. 2020)	Office of the Registrar General of India (2011)	9
Wealth status	The economically wealthier section can make better arrangements for rigid interventions such as lockdown while poor people may be forced to go out to arrange food supplies. Accessing meals from community kitchens, receiving food items from donations, buying essentials items from the government fair shops, may elevate the probability of the less economically well-off sections to catch the infection.	IIPS and ICF (2017)	8
Percentage of households exposed to smoking	Smoking may not only increase the risk of viral transmission but also raise the probability of dying, as the COVID-19 primarily afflicts the pulmonary system.	IIPS and ICF (2017)	8
Malarious districts/chloroquine drug distributed areas	Generally, malaria cases are calculated by Annual Parasite Index (API) [i.e. Plasmodium vivax (Pv) and Plasmodium falciparum (Pf)]. In malarious districts, chloroquine is the main drug used to alleviate symptoms. Since recent research proves that chloroquine has the potentiality to treat patients infected by SARS-CoV-2/COVID-19 (Colson et al. 2020; Gao et al. 2020), it is assumed that immunity capabilities of people in malarious districts shall be higher those in other areas	NVBDCP and Battle et al. 2015	9
Population per district or sub-district hospital	A higher population per health care institution implies lower resilience of an area. Although the PHCs/CHSs may not be capable enough to deal with COVID cases especially, they can aid in providing services for other morbidities, while the hospitals deal with Covid-19 cases.	MoHFW (2016); Office of the Registrar General of India (2011)	9

Source: Prepared by the authors

Appendix 1: Hotspot identification

The Getis-Ord local statistic is given as:

$$G_i^* = \frac{\sum_{j=1}^n w_{i,j} x_j - \bar{X} \sum_{j=1}^n w_{i,j}}{S \sqrt{\frac{n \sum_{j=1}^n w_{i,j}^2 - (\sum_{j=1}^n w_{i,j})^2}{n-1}}} \quad (\text{Eq. 1})$$

where x_j is the attribute value for feature j , $w_{i,j}$ is the spatial weight between feature i and j , n is equal to the total number of features. Furthermore,

$$\bar{X} = \frac{\sum_{j=1}^n x_j}{n} \quad (\text{Eq. 2})$$

and

$$S = \sqrt{\frac{\sum_{j=1}^n x_j^2}{n} - \bar{X}^2} \quad (\text{Eq. 3})$$

As the G^*i statistic is a z-score, no further calculations are required.

Appendix 2: COVID-19 potential outreach analysis

The COVID-19's potential outreach was assessed using the vector based double and vector based integer assessment through Python coding. The Euclidean distance was computed from each district's centroid and the number of COVID-19 cases for that district was used as its mass. As such, the higher the number of COVID-19 cases in a district, the higher would be its outreach potentiality towards the surrounding districts. This was computed as:

$$\text{EuclDistance} = \sqrt{(X_i - X_j)^2 + (Y_i - Y_j)^2} \quad (\text{Eq. 4})$$

$$\text{DIST PO}_{ij} = \frac{\text{No of COVID 19 cases for each dist}}{(\text{EuclDistance} * \text{RoadDistMajoration})^{\text{Impedance}}} \quad (\text{Eq. 5})$$

$$\text{UNITPO}_i = \text{UNITPotential}_i + \text{UNITPotential}_{ij} \quad (\text{Eq. 6})$$

where,

$i=1$ to total number of districts

$j=1$ to total number of districts

X_{ij} and Y_{ij} = coordinate from the district centre

Impedance (Impedance factor α) = 1

Road distance Majoration = 1.30

$\text{UNITPotential}_i = 0$

The estimated COVID-19 potential value was normalized in the range of 0-1 for weighting each conditioning parameter. Normalization was carried out, following Choi et al. (2009), as:

$$D_i = \frac{p_i - p_{\min}}{p_{\max} - p_{\min}} \quad (\text{Eq. 7})$$

If, p_i ($i = 1, 2, \dots, n$), then D_i would be the normalized district potential outreach value of p_i and p_{\min} and p_{\max} would denote the minimum and maximum normalized values of p_i , respectively. The normalized district potential outreach values were scaled between 0 to 1 and the natural breaks classification scheme was used to group the districts based on their respective values.

Appendix 3: Area-based susceptibility and vulnerability analysis: Statistical specifications

The susceptibility and vulnerability indices are the composite forms of the summation of normalized weighted variables. Before, beginning the statistical enumeration, all variables were converted to positive directions (i.e. a higher value would show a greater risk of susceptibility or vulnerability). The index construction procedure can be specified as follows:

$$Index = nm \sum_{n=1}^n (w * x_{nm}) \quad (\text{Eq. 8})$$

where, *Index* stands for the susceptibility or vulnerability index, *w* stands for the weight assigned to each indicator, and *x* and *n* denote the name and number of indicators respectively included in the analysis, while *nm* refers to the normalized function of the indicators and the sum of the normalized weighted indicators. Hence, x_{nm} denotes the normalized value of each indicator. In order to obtain scale-free indicators, the raw indicators were transformed by using the division by mean method (Kundu 1984, 2004 c.f. Haque 2016) as follows:

$$x_{nm} = \frac{x}{x_{mean}} \quad (\text{Eq. 9})$$

Simultaneously, the normalization of the sum of the normalized weighted indicators was computed as:

$$nm = \left(\frac{x - x_{minimum}}{x_{maximum} - x_{minimum}} \right) * 100 \quad (\text{Eq. 10})$$

The final output values obtained ranged between 0 and 100, with a higher value representing a greater risk of susceptibility or vulnerability to the COVID-19.

Appendix 4: Overall situation on 1st June, 2020, at the end of the various lockdown phases

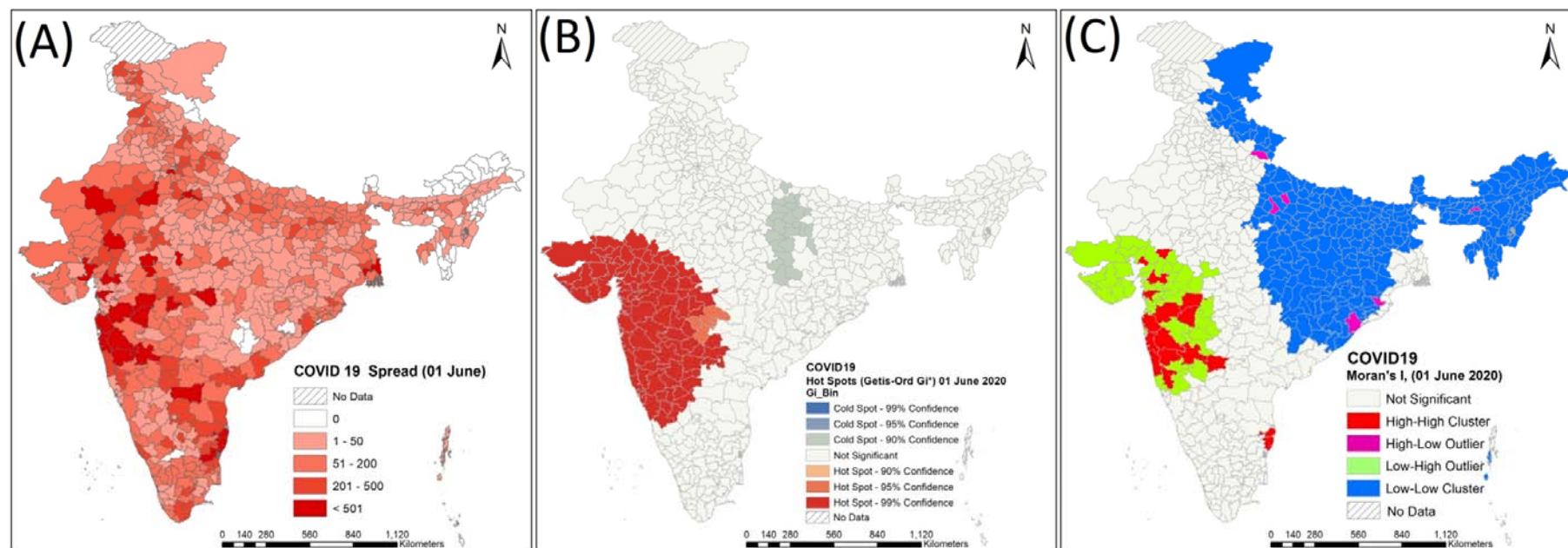


Figure A.1: Lockdown end status on 1st June, 2020, showing (A) District-wise spread of COVID-19 cases, (B) Hot/Cold Spots discerned, and (C) Significant clusters

It becomes apparent that the spread of COVID-19 in India could not be wholly attained simply by the various lockdown measures. This is evidenced by the overall continued rise in case numbers throughout the country [Figure A.1(A)], though crucially, the lockdown may have brought some relief in terms of a lowered rate of increase and the time to arrange for or create adequate medical care and quarantine facilities. The nodal growth areas are again around the principal metro cities, with particularly Mumbai, Delhi and Chennai reporting large numbers of cases. Notably, the virus has spread substantially in the north-eastern part of India, where its footprint was quite absent in the previous phases enumerated, and northern Kerala shows an upturn in cases as well. Greater mobility during the latter lockdown phases as a result of lowered restrictions and the carrier effects of migrants returning home have boosted case numbers throughout, hinting at the large-scale community transmission that is underway. Only in the central portion of the country (in parts of Madhya Pradesh and mostly Chhattisgarh) and in some districts of the North-east are cases still quite low. Hotspots are clustered mostly in the western region [Figure A.1(B)] and the rise of cases in the east has led to the decline of significant cold spots that existed previously in that region. Cases in the eastern part of the country are more uniformly spread while being more concentrated in certain districts in the west, as evident from the low-low and low-high clustering patterns [Figure A.1(C)].
