© 2025 IJNRD | Volume 10, Issue 5 May 2025 | ISSN: 2456-4184 | IJNRD.ORG

IJNRD.ORG ISSN : 2456-4184
INTERNATIONAL JOURNAL OF NOVEL RESEARCH AND DEVELOPMENT (IJNRD) | IJNRD.ORG
An International Open Access, Peer-reviewed, Refereed Journal

SPATIAL ANALYSIS OF COVID-19 SPREAD IN INDIA: INTEGRATING GIS AND AHP TO EVALUATE SOCIO-ENVIRONMENTAL RISK FACTORS

Ashutosh Mishra, Dr. Ashish Ratn Mishra, Dr. Mukesh Kumar, Prof. (Dr.) C. John Wesley, Prof. (Dr.) Deepak Lal

Associate Professor, Professor

Vaugh Institute of Agriculture Engineering & Technology, ShuatsCenter for Geospatial Technologies Vaugh Institute of Agriculture Engineering & Technology, Shuats

1. Introduction

The coronavirus disease 2019 (COVID-19) pandemic has escalated into an unprecedented global public health crisis (Wang, 2020). As of June 17, 2021, the virus had infected more than 176 million individuals across over 200 countries and regions, resulting in over 3.8 million confirmed deaths worldwide (https://www.worldometers). Beyond its immediate health impact, the pandemic has imposed a devastating economic burden and has strained healthcare systems globally. The continued rise in infections and fatalities underscores the evolving and severe nature of the crisis.

In India, COVID-19 has spread extensively, affecting both urban and rural populations. However, the intensity of transmission and mortality rates have varied significantly across regions. Some areas experienced widespread outbreaks, while others reported relatively lower-case numbers. These disparities are likely influenced by a combination of factors, including population density, healthcare infrastructure, mobility patterns, public health interventions, and socio-economic conditions. Despite extensive efforts, the precise determinants of disease spread and the variation in mortality rates remain only partially understood, reflecting the complex epidemiology of the virus. Many studies have previously examined the causes of COVID-19 (Kramer, 2006; Kampf 2020; Rothe,2020; Kupferschmidt,2020; Bai, 2020).

In this context, geospatial technologies—particularly Geographic Information Systems (GIS)—have gained increasing prominence in the field of public health. Over the past several years, GIS has proven to be a vital tool for infectious disease surveillance, spatial modeling, and response planning. The COVID-19 pandemic has further highlighted the need for real-time mapping, data integration, and spatial analytics to support decision-making and resource allocation. As a result, there has been a marked increase in the adoption of information-age technologies for monitoring communicable diseases, guiding public health interventions, and enhancing epidemiological research. Many researchers have also conducted GIS-based studies related to COVID-19 (Agrawal et al., 2022; Urban et al., 2021; Scarpone et al., 2020; Schmidt et al., 2021; Praharaj et al., 2022).

In this study, several key parameters were considered to assess their relative influence on the spread and severity of COVID-19. These parameters include the Air Quality Index (AQI), the proportion of the population aged over 65 years, population density, literacy rate, and the availability of medical professionals (doctors). To systematically evaluate the importance of each factor, the Analytic Hierarchy Process (AHP) was employed. This multi-criterion decision-making technique allows for the comparison of these variables in a structured manner, helping to identify which factors exert the most significant impact on public health outcomes during the pandemic.

2. Material and Method

Study area

India, the second-most populous country in the world, presents a highly diverse demographic, socio-economic, and environmental landscape, making it an ideal case study for spatial health analytics. This study encompasses all **28 states** and **8 union territories** of India, covering a broad geographic expanse and a wide range of public health conditions, urbanization patterns, population densities, and environmental indicators. The heterogeneity of these factors provides a meaningful foundation for analyzing how spatial determinants influence the spread and severity of infectious diseases such as COVID-19.

India experienced multiple waves of COVID-19 between 2019 and 2021, with varying intensity across regions. Some areas, particularly densely populated urban centers like Maharashtra, Delhi, and Tamil Nadu, recorded high infection rates and fatalities, while others with more rural or dispersed populations, such as the northeastern states, showed comparatively lower transmission rates. This variation underscores the influence of spatial and socio-environmental determinants, which this study aims to quantify and visualize.

To understand these dynamics, the study focuses on five critical factors across each state and union territory: **Air Quality Index (AQI)**: Air pollution levels, which can aggravate respiratory illnesses and increase susceptibility to COVID-19. **Age-Wise Population (AWP)**: The percentage of population aged 65 years and above, who are more vulnerable to severe disease outcomes. **Population Density (PD)**: A key driver of virus transmission due to close human contact. **Literacy Rate (LTR)**: Reflective of public awareness and the capacity to comprehend and comply with health advisories. **Availability of Doctors (ABD)**: Indicative of healthcare infrastructure strength and the capacity to manage pandemic response. Each state's data for these parameters were gathered from reliable government sources, such as the Census of India, the Ministry of

Health and Family Welfare, and the Central Pollution Control Board. COVID-19 infection data from the MyGov.in served as the dependent variable for the regression models.

By applying spatial interpolation and statistical modelling across this diverse national landscape, the study aims to generate interpolated maps that reflect regional disparities in vulnerability and impact. These maps provide policymakers with a clearer picture of spatial risk, enabling targeted interventions and improved pandemic preparedness at both local and national levels. The selected study area thus not only serves the research objectives of understanding COVID-19 spread through a spatial lens but also demonstrates the adaptability of the proposed Web-GIS platform to large, complex, and heterogeneous geographies like India.

Data Used

This study examines the impact of various factors on Covid-19, including Air Quality Index (AQI), the population aged over 65 years (AWP), Population Density (PD), Literacy Rate (LTR), and the Availability of Doctors (ABD). Covid-19 data for all 28 states and 8 union territories of India (2019-2021) were used. AQI data was sourced from 20 states for 2019, while Age-wise Population (AWP), Population Density (PD), Literacy Rate (LTR), and Availability of Doctors (ABD) data were obtained from the 2011 census for all 28 states and 8 union territories. The detail of datasets collected from different source are given in Table 1.

Table 1: Acquired datasets and their sources

Parameters	Source
Covid-19	https://www.mygov.in/covid-19-archive
AQI	https://www.data.gov.in/resource/stateut-wise-number-days-different-categories-air-quality-index-aqi-cities-country-2019
AWP	https://censusindia.gov.in/nada/index.php/catalog/1541
PD	https://www.indiacensus.net/density.php
LTR	https://www.indiacensus.net/literacy-rate.php
ABD	(Lok sabha unstarred question no. 2067 to be answered on 02nd august, 2024) https://sansad.in/getFile/loksabhaquestions/annex/182/AU2067_PZDQxt.pdf?source=pqals

3. Methodology

Various factors have been considered and assigned weights to evaluate the relative impact of different parameters on the spread and severity of COVID-19.

Assessment of weights normalized using the Analytic Hierarchy Process (AHP)

In the present study Analytical Hierarchy Process (AHP) developed by Saaty (1980) was used as a decision maker for computing the weights which were assigned to various parameters and the value of its impact on Covid-19. AHP is used for judgments of the problem into a hierarchy of the criteria and its importance is given based on the topic related to the study. Hierarchy is very important as it helps in decision making by evaluating the problem separately of each parameter (Saaty 1990). Paired wise comparison matrix was made to compare all the parameters and how it's related by giving the value of its importance by using Saaty's (1980) method, a scale of 1-9 to determine the relative importance values of all the parameters.

The weight of various parameters was decided based on the result of the multiple linear regression, for the study to determine the impact of these parameters on Covid-19. The weights were assigned to different parameters by using a pairwise comparison matrix and then each feature of the parameters was normalized using Saaty's AHP to reduce the subjectivity associated with the assigned weights. In each pair of the criteria, the decision maker is asked to which degree a criterion is more important than the other. After comparisons, the method defines the relative position of one criterion in relation to all other criteria. The Saaty's method used a nine-point scale which is assigned to the criteria as per value used for analysis. Then the normalized weight of each criterion is derived by calculating the total of each feature of the criteria divided by the number of criteria used for the study. The consistency vector of the diagonal value of each criterion is calculated using the Eigen value matrix technique.

The normalization process converts the measurements of a set of objects on a standard scale into relative scale measurements (Saaty, 1990). Saaty (1980) suggested that the weights assigned and the normalized weights of the parameters and each feature were examined for consistency weights.

Table 2: Scale for pair-wise comparison matrix

Intensity Importance	Linguistic variables
1	Equal importance
2	Equal to moderate importance
3	Moderate importance
4	Moderate to the strong importance
5	Strong importance
6	Strong to the very strong importance
7	Very strong importance
8	Very to the extremely strong importance
9	Extreme importance

The following steps were carried out to compute the final weights of all the parameters:

1. Sum the values in each column of the pair-wise comparison matrix using the formula,

$$L_{ij} = \sum_{n=1}^{n} C_{ij} \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \quad [1]$$

Where L_{ij} is the total column value of the pair-wise comparison matrix and C_{ij} are the criteria used for the analysis, i.e., AQI, population density, ABD etc.

2. Divide each element in the matrix by its total row to generate a normalized pair-wise comparison matrix.

$X_{ij}=\frac{C_{ij}}{\sum_{n=1}^{n} C_{ij}}$.. [2]

Where X_{ij} = normalized pair-wise comparison matrix

3. Divide the sum of the normalized row of the matrix by the number of criteria/parameter (N) to generate the standard weight by using the following formula,

$W_{ij} = \frac{\sum_{j=1}^{n} X_{ij}}{N}$... [3]

WhereW_{ij} = Standard weight

4. For calculating the consistency vector values the following formula was used:

$\lambda = \sum_{i=1}^{n} CV_{ij}$, Where λ= Consistency vector ………….. [4]

5. Consistency Index (CI) was used as a deviation or degree of consistency which was then calculated using the formula below:

CI $=\frac{\lambda-n}{n-1}$, Where CI = Consistency Index, n = Number of criteria …….. [5]

6. Consistency ratio (Cr) is calculated by using the formula:

Cr $= \frac{CI}{RI}$……………….. [6]

Where, RI = random inconsistency

If the value of Consistency ratio is less than or equal to 0.10 then the inconsistency is acceptable. Random inconsistency values for ‘n’ number of criteria i.e. number of parameters.

Table 3: Random inconsistency values (Saaty, 1980)

n	2	3	4	5	6	7	8	9
RI	0	0.52	0.9	1.12	1.24	1.32	1.41	1.45

Where n = number of criteria used and RI = Random Inconsistency

4. Result and Discussion

4.1. Assignments of weights and weight normalization of the Five factors affecting COVID-19

An assessment of No. of cases of COVID-19 in India was performed using five different factors including Air Quality Index (AQI), the population aged over 65 years (AWP), Population Density (PD), Literacy Rate (LTR), and the Availability of Doctors (ABD). AHP model and QGIS software were used for generating, integrating and analysing the data utilized in the present study. The assignments of weights and weights normalization of the five factors are presented below:

Based on study and expert 's ranking a pair wise comparison of the five factors were computed in a square matrix, where diagonal elements of the matrix are always 1 (Table 4). All the thematic layers were compared against each other to get the normalized weights of each theme. Likewise, each theme was compared with other theme and the matrix was developed which is shown in the Table 4. The normalized pair wise matrix was calculated by using equation [2] which is shown in Table 5. Final normalized weight was obtained by using equation [3] and shown in Table …………….

Table 4 Pair-wise comparison matrix of five factors

	AQI	AGE	ABD	PD	LTR
AQI	1	5	9	7	7
AWP	0.2	1	6	5	4
ABD	0.111111111	0.166667	1	6	4
PD	0.142857143	0.2	0.166667	1	9
LTR	0.142857143	0.25	0.25	0.111111	1
TOTAL	**1.596825397**	**6.616667**	**16.41667**	**19.11111**	**25**

Table 5 Normalized Pair-wise Matrix

	AQI	AWP	ABD	PD	LTR	TOTAL	NOR. WT.
AQI	0.626243	0.755668	0.548223	0.366279	0.28	2.576412	0.52
AWP	0.125249	0.151134	0.365482	0.261628	0.16	1.063492	0.21
ABD	0.069583	0.025189	0.060914	0.313953	0.16	0.629639	0.13
PD	0.089463	0.030227	0.010152	0.052326	0.36	0.542168	0.11
LTR	0.089463	0.037783	0.015228	0.005814	0.04	0.188289	0.03

Table 6 Normalized Weight of Five factors

Factors	AQI	AWP	PD	ABD	LTR
Normalized Weight	0.52	0.21	0.13	0.11	0.03

Consistency Analysis:

Consistency vector is calculated by multiplying the pairwise comparison matrix values and normalized weights of the five factors accordingly using matrix multiplication which is shown in table

Table 7 Consistency Analysis

	AQI	AGE	PD	DA	LITER
AQI	**3.0589949**	**2.2345679**	**3.14750542**	**4.52173913**	**4.7619**
AWP	**1.35906773**	**0.74074074**	**0.96746204**	**2.51304348**	**2.5905**
PD	**1.20757465**	**0.61111111**	**0.32537961**	**0.86413043**	**1.3591**
ABD	**0.52221413**	**0.3382716**	**0.27548807**	**0.32608696**	**0.4619**
LTR	**0.2521243**	**0.20047031**	**0.19801673**	**0.20263975**	**0.2381**

Then, the consistency vector (λ) was calculated using equation [4] which is found to be 4.689 (3.059+0.741+0.325+0.326+0.238). Further using equation [5] and [6] the value of consistency index (CI) and consistency ratio (Cr) was calculated as -0.09 and -0.08 respectively. The value of consistency ratio (Cr) obtained is less than 0.1, hence the inconsistency is acceptable (Muralitharan and Palanivel, 2015).

4.2. State wise Assignments of weights and weight normalization of the Five factors affecting COVID-19

The area-specific AHP weights of five key socio-environmental factors—Air Quality Index (AQI), Age-wise Population (AWP), Population Density (PD), Literacy Rate (LTR), and Availability of Doctors (ABD) was calculated for each state and union territory in India. The AHP (Analytic Hierarchy Process) was used to determine the relative importance of each factor at the regional level, reflecting how localized variations influence COVID-19 vulnerability. The matrix was then normalized to compute a set of weights, where each value represents the proportional influence of a specific factor in that area. This area-wise breakdown of factor influence enables targeted interpretation. For example, if AQI weight is high in Delhi, it indicates that air quality plays a dominant role in COVID-19 spread in that region. Similarly, a higher AWP weight in Kerala reflects the stronger impact of the elderly population on infection vulnerability there.

4.3. State wise Assignments of weights and weight normalization depicting the combined impact of the Five factors affecting COVID-19

The overall impact score of each area was calculated as a composite weighted sum of all five factors. The individual weights of each factor per area was used to derive a single normalized score indicating the relative overall vulnerability or severity of COVID-19 in that region.

The composite weight represents a normalized weight that reflects the combined influence of AQI, AWP, PD, LTR, and ABD for that particular state or union territory. The result allows for spatial comparison across the country, helping to identify: High-impact regions (e.g., states with high composite weights like Maharashtra or Uttar Pradesh), Moderate-risk regions, and Low-risk regions, based on the normalized values.

4.4. Interpolated Heatmaps

The localized factor weights were later used to perform factor-wise interpolations, helping visualize the spatial distribution of individual determinant severity across India whereas the composite weights were used for generating the final interpolated map showing the cumulative spatial impact of all the five factors on

COVID-19 across India. This visualization provides a critical decision-support layer for policymakers to prioritize resource allocation, public health interventions, and strategic planning.

Fig 1 Interpolated map for the factor ABD

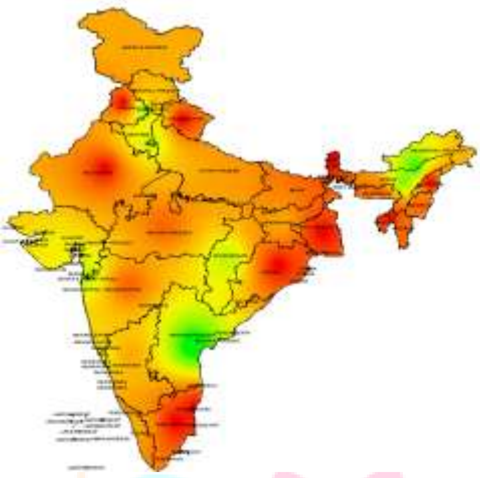

Fig 2 Interpolated map for the factor AWP

Fig 3 Interpolated map for the factor AQI

Fig 4 Interpolated map for the factor PD

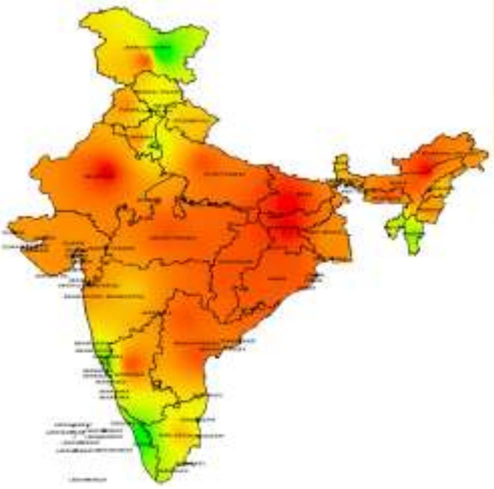

Fig 5 Interpolated map for the factor AQI

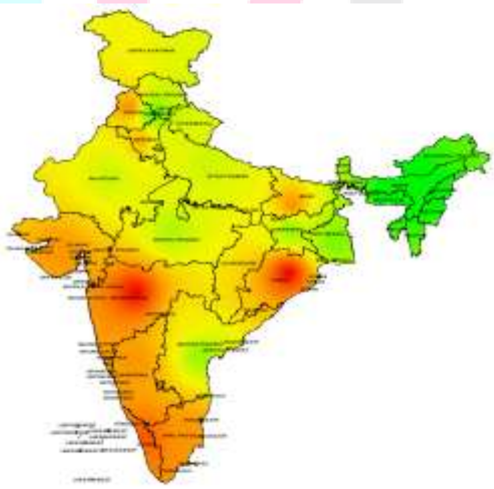

Fig 6 Interpolated composite map for all the factors

Legend: Disease Impact Intensity Across Indian States

Color	Interpretation
Red	**Very High Impact** (Severe cases or high vulnerability)
Orange	**High Impact** (Elevated risk or spread)
Yellow	**Moderate Impact** (Noticeable but controlled)
Green	**Low/Negligible Impact** (Minimal or well-contained)

5. Conclusion

By integrating geospatial technologies, statistical modeling, and multi-criteria decision analysis (MCDA), the research offers an innovative method for assessing and visualizing the spatial impact of various socio-environmental determinants on disease spread and severity. The analysis focused on five key factors—Air Quality Index (AQI), the proportion of the elderly population (AWP), population density (PD), literacy rate (LTR), and availability of doctors (ABD)—across all 28 states and 8 union territories of India. Linear and multiple regression models were employed to quantify the individual and combined influence of these factors on COVID-19 infection rates. The regression outputs were then integrated into the Analytic Hierarchy Process (AHP) framework to calculate both global and localized weights, allowing for a nuanced understanding of factor importance in different regions. Spatial interpolation techniques, including inverse distance weighting (IDW), were applied to generate continuous surface maps for each factor and for the combined impact of all factors. These maps provide a powerful visual representation of spatial health risks, revealing both patterns and outliers in disease vulnerability. This study may be helpful in developing a Web-GIS platform using Django, Plotly, and Leaflet.js offering a dynamic, user-friendly interface that supports uploading custom datasets, selecting variables, generating regression outputs, calculating AHP weights, and visualizing spatial data in future. Notably, the system is adaptable for analysing diseases beyond COVID-19, enabling users to input their own parameters and datasets to study other health outcomes. In conclusion, this research not only enhances our understanding of COVID-19 through spatial analysis but also establishes a robust, generalizable framework for public health decision-making. The integration of statistical analysis, MCDA, and geospatial visualization in a web-based platform would mark a significant step toward real-time, data-driven epidemiological intelligence in near future. The findings and tools developed in this study hold strong potential for informing targeted health interventions, optimizing resource allocation, and improving pandemic preparedness and response strategies in the future.

References

1. Agrawal, S., & Agrawal, S. (2022). Spatial mapping and cluster analysis of COVID-19: A case study of Uttar Pradesh, India. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, *X-4/W3-2022*, 3–10. https://doi.org/10.5194/isprs-annals-X-4-W3-2022-3-2022

2. Bai, Y., Yao, L., Wei, T., Tian, F., Jin, D. Y., Chen, L., et al. (2020). Presumed asymptomatic carrier transmission of COVID-19. *JAMA*. https://doi.org/10.1001/jama.2020.2565

3. Worldometers. (2020). Coronavirus outbreak. https://www.worldometers.info/coronavirus/ (Accessed February 23, 2020)

4. Kampf, G., Todt, D., Pfaender, S., & Steinmann, E. (2020). Persistence of coronaviruses on inanimate surfaces and their inactivation with biocidal agents. *Journal of Hospital Infection*. https://doi.org/10.1016/j.jhin.2020.01.022

5. Kramer, A., Schwebke, I., & Kampf, G. (2006). How long do nosocomial pathogens persist on inanimate surfaces? A systematic review. *BMC Infectious Diseases*, *6*, 130. https://doi.org/10.1186/1471-2334-6-130

6. Kupferschmidt, K. (2020). Study claiming new coronavirus can be transmitted by people without symptoms was flawed. *Science*. https://www.science.org/content/article/study-claiming-new-coronavirus-can-be-transmitted-people-without-symptoms-was-flawed

7. Praharaj, S., Solis, P., & Wentz, E. A. (2022). Deploying geospatial visualization dashboards to combat the socioeconomic impacts of COVID-19. *Environment and Planning B: Urban Analytics and City Science*, *50*(5), 1262–1279. https://doi.org/10.1177/23998083221142863

8. Rothe, C., Schunk, M., Sothmann, P., Bretzel, G., Froeschl, G., Wallrauch, C., et al. (2020). Transmission of 2019-nCoV infection from an asymptomatic contact in Germany. *New England Journal of Medicine*. https://doi.org/10.1056/NEJMc2001468

9. Scarpone, C., Brinkmann, S. T., Große, T., Sonnenwald, D., Fuchs, M., & Walker, M. (2020). A multimethod approach for county-scale geospatial analysis of emerging infectious diseases: A cross-sectional case study of COVID-19 incidence in Germany. *International Journal of Health Geographics*, *19*, 32. https://doi.org/10.1186/s12942-020-00226-0

10. Schmidt, F., Dröge-Rothaar, A., & Rienow, A. (2021). Development of a Web GIS for small-scale detection and analysis of COVID-19 (SARS-CoV-2) cases based on volunteered geographic information for the city of Cologne, Germany, in July/August 2020. *International Journal of Health Geographics*, *20*, 40. https://doi.org/10.1186/s12942-021-00290-0

11. Urban, R. C., & Nakada, L. Y. K. (2021). GIS-based spatial modelling of COVID-19 death incidence in São Paulo, Brazil. *Environment and Urbanization*, *33*(1), 229–238. https://doi.org/10.1177/0956247820963962

12. Wang, C., Horby, P. W., Hayden, F. G., & Gao, G. F. (2020). A novel coronavirus outbreak of global health concern. *The Lancet*. https://doi.org/10.1016/S0140-6736(20)30185-9

13. Murugesan, V. S., Suresh, S., & Prabakaran, M. (2019). Enhancement of mail operational performance of India Post facility layout using AHP. *Journal of Engineering Research and Application*, *9*(9), 31–37.