Statistica & Applicazioni - Vol. XXI, n. 1, 2023, pp. 66-79
Published online march 2026

RURAL HEALTHCARE INFRASTRUCTURE IN INDIA: A COMPARATIVE ANALYSIS ACROSS PRE-PEAK-AND POST-COVID-19 PERIODS

Sindhu Marangattu Sindhumo*
Shanmugam Bavadhaarini*
Tullio Menini**

SUMMARY

The COVID-19 pandemic has posed unprecedented challenges to healthcare systems worldwide, with rural healthcare infrastructure in India facing significant strain. This study undertakes a time-phased analysis of the development of rural healthcare infrastructure in India, focusing on the pre-COVID (2019-20), peak-COVID (2020-21), and post-COVID (2021-22) periods. The research assesses key indicators such as healthcare accessibility, resource allocation, workforce challenges, and the adaptability of healthcare services. The findings highlight the initial disruptions caused by lockdowns, subsequent efforts to bolster healthcare resources, and the long-term effects on healthcare delivery in rural areas by using spatial analysis. Additionally, the study discusses the role of government policies and community responses in mitigating the pandemic's impact. This comprehensive analysis informs policymakers and stakeholders about the critical gaps and potential strategies for strengthening rural healthcare infrastructure to withstand future public health crises.

Keywords: Healthcare Infrastructure, COVID-19 Pandemic, Principal Component Factor Analysis (PCFA), Sustainable Development Goals (SDG), Spatial Analysis.
DOI: 10.26350/999999_000059
ISSN: 1824-6672 (print) 2283-6659 (digital)

1. INTRODUCTION

Healthcare in rural areas is symbolic of a nation's commitment to equitable well-being. The glaring disparities between urban and rural healthcare in India have persisted, posing substantial challenges to comprehensive healthcare accessibility. Rural healthcare infrastructure encompasses the physical, organizational, and human resources necessary to deliver healthcare services to rural populations (Goel, 2011). It includes a network of primary healthcare centers, sub-centers, and community health centers, as well as the healthcare professionals, medical equipment, and essential facilities required to provide medical care to rural residents. As per the population norms, one Sub-center is established for every 5000 population in plain areas and for every 3000 population in hilly/tribal/desert areas. There were

* Department of Statistics - University of Madras - Chennai-600004 - TAMILNADU, India (email: mrsindhu@gmail.com; sindhu@unom.ac.in).
** Department of Human and Social Sciences - University of Naples "L'Orientale" - Largo S. Giovanni Maggiore, 30 - 80134 NAPLES (✉ email: menini@unior.it).

161829 Sub-centers in India as of March 31, 2022- 157935 of which functioned in rural areas.

For the development of the health status of the citizens of the country there is a need for adequate health care infrastructure. In February 2018, the Government of India announced the creation of 150000 Health and Wellness Centers (HWCs) by transforming the existing Sub Centers and Primary Health Centers. These centers are to deliver Comprehensive Primary Health Care (CPHC) bringing healthcare closer to the homes of people. They cover both, maternal and child health services and non-communicable diseases, including free essential drugs and diagnostic services.

The Sustainable Development Goals (SDGs), adopted by the UN in 2015, encompass 17 global objectives to tackle social, economic, and environmental issues.

The research directly addresses the core objective of SDG 3 by evaluating the accessibility, availability, and quality of healthcare services in rural India, thereby assessing progress toward achieving universal health coverage and improved health outcomes for all. The COVID-19 pandemic exposed significant weaknesses in rural health systems, including shortages of medical personnel, inadequate facilities, lack of insurance coverage and limited emergency preparedness, which hindered the realization of equitable healthcare delivery. By comparing pre-, peak-, and post-pandemic conditions, the study provides insights into how health infrastructure resilience and policy responses have influenced the attainment of SDG 3 targets. Simultaneously, the research engages with SDG 10, which focuses on reducing inequalities within and among populations. The urban-rural divide in healthcare access, quality, and resource distribution represents a major structural inequality in India's health system. Examining these disparities over different phases of the pandemic highlights whether policy interventions and investments have contributed to narrowing or widening these gaps. Overall, the study focuses on ensuring healthy lives and well-being for all and addressing health disparities to provide equitable healthcare access for vulnerable populations.

Hence this article embarks on an intricate exploration of the spatial distribution of rural healthcare infrastructure across all 28 States, which have separate state governance under Central governance and 8 Union Territories (UTs) which have direct Central governance in India. The study excluded Chandigarh UT due to the lack of data, delving into the critical spatial dimensions shaping rural healthcare development during pivotal epochs. Harnessing the wealth of available data, the study aims to unveil the complex dynamics of rural healthcare, guided by the principles of equality and equity. While equality offers the same resources and opportunities to all, equity acknowledges diverse needs and circumstances, ensuring access to necessary care regardless of socioeconomic status or location. This involves targeted interventions and resource allocation to support marginalized and underserved communities, and it reduces the healthcare outcome gap between rural and urban areas. By delving into the spatial dimensions, the research aims to inform policies that address immediate challenges while embracing the unique characteristics of different regions, thus fostering more equitable healthcare access and development.

The structure of the paper is outlined as follows: in Section 2 we review existing literature on rural health disparities and identify the key research gaps. Section 3 describes the data sources. Section 4 deals with the methodological framework, detailing the construction of the Healthcare Infrastructure Index (HII) and the application of Principal Component Factor Analysis (PCFA). In particular, spatial analysis using Geographic Information System (GIS) tools to visualize regional inequalities is presented in Section 5. Results and interpretation of temporal and regional variations are discussed in Section 6, while Section 7 provides concluding remarks in the context of SDG 3 (Good Health and Well-being) and SDG 10 (Reduced Inequalities).

2. REVIEW OF LITERATURE

The literature review centers on five seminal studies that explore methodologies for calculating the HII and conducting spatial analysis across Indian States and UTs.

Bezerra, Santos, Lisbinski and Dias (2020) detailed the methods for constructing the HII in Brazil, emphasizing its spatial distribution. Their research highlights Brazil's health infrastructure challenges, particularly during COVID-19, noting the unequal distribution of primary and specialty care, regional disparities in care quality, and the shortage of Intensive Care Unit (ICU) beds in the public health sector compared to the private health sector.

Abdi (2003) emphasized the importance of factor rotation in factor analysis to enhance interpretability. After extracting and retaining key factors, they undergo rotation either orthogonal or oblique to align them meaningfully. While rotated axes may explain less individual variance, the total variance remains constant, aiding interpretation over statistical optimization. This concept, often illustrated using Principal Component Analysis (PCA), and subject scores, with careful subspace size selection crucial for robust interpretation.

Gere (2023) highlighted Euclidean distance and Ward's method as reliable clustering techniques due to their intuitive applicability and effectiveness in forming cohesive, distinct clusters. Euclidean distance measures straight-line proximity in multidimensional space, while Ward's method minimizes intra-cluster variance and handles outliers robustly, making it suitable for noisy datasets.

Tsai, Lin, Chu and Perng (2009) employed spatial autocorrelation methods, including Global Moran's I and Local Getis-Ord, to map disease clusters and analyze spatial patterns for the leading causes of death in Taiwan. This analysis aids in understanding spatial risk factors and informs healthcare policy development.

Mathur (2015) categorized spatial autocorrelation techniques in ecology into global, local, and variogram measures. Global methods, like Moran's I, assess overall spatial clustering, whereas local measures, such as Local Indicators of Spatial Association (LISA) (Anselin, 1995), identify specific hotspots. Despite Moran's I's prevalence, its application in Indian plant population studies is limited, presenting future research opportunities.

This study aims to assess India's health infrastructure by estimating its HII. The study also provides insights into the methodologies employed and their implications for healthcare policy and practice in India, particularly within the spatial context.

3. DATA SOURCE

The rural health infrastructure for the 28 States and 8 UTs for 3 years Pre-COVID (2019-2020), Peak-COVID (2020-2021), and Post-COVID (2021-2022) has been collected from the Rural Health Statistics Report in Health Management Information System (HMIS) Portal (Rural Health Statistics 2018-19, Government of India Ministry of Health and Family Welfare Statistics Division), Table 1. This data serves as a critical resource for informed decision-making and strategic planning aimed at enhancing healthcare delivery in rural areas. The selection of study during COVID period enables us to identify the strengths, weaknesses, and areas for improvement in rural healthcare infrastructure toughness and preparedness during public health emergencies (Anand, 2014).

TABLE 1. - *Healthcare Variables and their Dimension*

Dimension	Variables
Staffing facilities	Number of Doctors
	Number of Pharmacists
	Number of Laboratory Technicians
	Number of Nursing Staff
	Number of Radiographers
	Number of Pediatrician
	Number of Health workers Male
	Number of Health workers Female
	Number of Total Specialists
	Number of Health Assistants
	Number of General Duty Medical Officers
Health facilities	Number of Toilets
	Number of Beds

As per the National Institute of Health study on 2019-2020, Maharashtra, Delhi, Gujarat, Tamil Nadu, Madhya Pradesh, Rajasthan, Uttar Pradesh, Kerala were states highly affected by the pandemic. The States and UTs of the data have been grouped into 4 regions namely the Eastern region which consists of Bihar, Jharkhand, Sikkim, Odisha, Manipur, Assam (Paul, Jana and Maiti, 2019). Nagaland, Mizoram, Tripura, Arunachal Pradesh, Meghalaya, Chhattisgarh, and West Bengal, the Western region which consist of Gujarat, Maharashtra, Madhya Pradesh, Goa, Daman and Diu and Dadra and Nagar Haveli, the Northern region Ladakh, Jammu and Kashmir, Punjab, Haryana, Himachal Pradesh, Rajasthan, Chandigarh, Uttarakhand, Delhi and Uttar Pradesh, and the Southern region Tamil Nadu, Kerala, Karnataka, Andhra Pradesh, Telangana, Puducherry, Andaman and Nicobar, and Lakshadweep. These regions have been divided through this DCMSME Ministry - Development Commissioner Ministry of Micro, Small and Medium Enterprises (Rajeeb, 2015).

Data transformation has been done through standardization following the procedure outlined by Fávero and Belfiore (2019). This process ensures that the variables are on the same scale, making them directly comparable and facilitating statistical analysis. The objective of this study is to analyze the development of Rural Healthcare Infrastructure for 3 different periods, say Pre- Peak- Post- COVID in States and UTs of India based on the calculated index namely HII. And study also investigates whether there is any progressive growth shown by States/UT regarding SDG 3 and SDG 10.

4. PROPOSED METHODOLOGY

4.1 *Healthcare Infrastructure Index (HII)*

The HII calculates a score for each state based on various indicators related to healthcare infrastructure. These indicators may include the availability of healthcare facilities, accessibility of services, quality of care, integration of technology, and allocation of resources. By aggregating these indicators into a single score for each state, the HII provides a comparative assessment of healthcare infrastructure across different states, helping policymakers identify areas for improvement and prioritize resource allocation to enhance healthcare delivery and outcomes. This HII is derived using PCFA along with the Kaiser-Meyer-Olkin (KMO) test and Bartlett’s test. These statistical techniques aid in refining the HII calculation process.

4.2 *Principal Component Factor Analysis (PCFA)*

The PCFA is the combination of the principles of PCA and Factor Analysis (FA) and it combines varimax rotation which is used to analyze the multidimensional data Trendafilov and Gallo (2021a,b). It allows us to identify underlying factors contributing to healthcare infrastructure quality by reducing the dimensionality of the dataset and extracting common factors. Subsequently, Varimax rotation is applied to enhance the interpretability of the factor solution by maximizing the variance of factor loadings. The KMO test assesses the adequacy of the data for factor analysis, while Bartlett's test examines whether the variables in the dataset are interrelated, thus validating the suitability of the dataset for PCFA.

$$HII_m = \sum_{j=1}^{p} \frac{\sigma_j^2}{\sum_{j=1}^{p} \sigma_j^2} F_{jm} \tag{1}$$

where HII_m is a the HII of the ith state; σ_j^2 is the variance explained by factor j; p is the number of selected factors; $\sum_{j=1}^{p} \sigma_j^2$ is the j sum of the variances explained by the extracted p factors; and F_{jm} is the factor score of the state m for the factor j.

An adequacy test, in the context of statistical analysis, refers to a diagnostic procedure performed to assess whether a particular statistical technique or model is appropriate for the dataset. The purpose of adequacy testing is to ensure that the assumptions underlying the chosen statistical method are met and that the results obtained from the analysis are reliable and meaningful. The KMO test is a statistical measure used to assess the sampling adequacy for factor analysis, including PCFA. It evaluates whether the variables in the dataset are suitable for factor analysis by examining the strength of correlation between variables. The KMO test produces a KMO statistic, which is a measure of sampling adequacy ranging from 0 to 1. A higher KMO value indicates that the dataset is more suitable for factor analysis, while a lower value suggests that the dataset may not be appropriate for this analysis. In general, when the KMO value is sufficiently high (typically above 0.6 or 0.7), it indicates that the dataset has enough correlation. The KMO test was conducted to assess the sampling adequacy for factor analysis in our dataset.

Bartlett's test is a statistical test used to determine whether the correlation matrix of the variables in a dataset is significantly different from the identity matrix, indicating whether the variables are interrelated and suitable for factor analysis. In essence, Bartlett's test assesses the null hypothesis that the variables in the dataset are uncorrelated, suggesting that factor analysis may not be appropriate. If the probability value of the test is less than level of significance, the variables are sufficiently correlated to justify conducting factor analysis. Bartlett's test complements the KMO test in assessing the adequacy of the dataset for factor analysis.

The HII score is standardized such that the maximum and minimum values are between 0 and 1. The closer the value is to 1, the better the COVID-19 response healthcare infrastructure in the respective state.

The PCFA based HII ensures a robust and reliable assessment of healthcare infrastructure across various regions. Further to show how the indices were spread across the country a spatial analysis is performed.

5. SPATIAL ANALYSIS

Spatial analysis examines attributes, locations, and relationships within spatial data using analytics, computational models, and algorithms to extract valuable insights. It goes beyond simple mapping by visualizing the interactions between variables, offering decision-makers a comprehensive perspective through consolidated information from multiple sources (Fávero and Belfiore, 2019). Identifying spatial patterns is crucial in spatial data analysis (Gatrell and Bailey, 1996), often tested using Moran's I statistic, which assesses spatial autocorrelation. If the null hypothesis of zero autocorrelation is rejected, it indicates spatial dependence.

Traditional global analyses like nearest neighborhood analysis, and the semivariogram are commonly used to study these patterns (Anselin and Rey, 2014).

In spatial analysis, a weight matrix represents the spatial relationships between observations in a dataset, with each element indicating the strength of the relationship based on proximity or similarity. Various techniques, such as spatial autocorrelation measures, spatial regression models, and spatial interpolation methods, require a weight matrix to analyze spatial patterns and processes. Two common types of weight matrices are contiguity-based and distance-based.

Contiguity-based weight matrices represent spatial relationships based on adjacency. In Queen's contiguity, observations are considered neighbors if they share a boundary or vertex, including diagonal connections. Conversely, in Rook's contiguity, observations are deemed neighbors only if they share a boundary, excluding diagonal connections.

Distance-based weight matrices represent spatial relationships based on the distances between observations. The Inverse distance weights method assigns higher weights to closer observations, inversely proportional to their distances. The Kernel weights method uses a kernel function to assign higher weights to closer observations and lower weights to those farther away. The K-nearest neighbors (KNN) method identifies the K-closest observations to a focal point using distance metrics like Euclidean or Manhattan distance. The parameter K determines the number of neighbors considered, influencing the granularity of the spatial analysis. Weights are typically assigned inversely to distances, reflecting the degree of influence or similarity of the neighbors on the focal observation, with closer neighbors receiving higher weights. The concepts of weight matrix creation, KNN method of weight matrix, and visual representation of the results were done through spatial statistical software QGIS 3.34.3 'Prizren' and GeoDa version 1.22.

6. ANALYSIS AND RESULTS

In this section, healthcare infrastructure dynamics during different phases of the COVID-19 pandemic is explored. Using statistical methods like PCFA in SPSS, the dataset is examined to uncover underlying patterns. With GIS tools such as QGIS, the spatial distribution of key indices and clusters, highlighting disparities in healthcare infrastructure, are mapped. Additionally, using GeoDa software, the clusters with similar healthcare infrastructure levels are identified. Through this analysis, a better understanding of healthcare resilience amidst global health challenges are possible.

TABLE 2. - *Results of PCFA during Pre, Peak and Post_COVID*

Adequacy Test	Pre_COVID	Peak_COVID	Post_COVID
KMO Test	0.719	0.756	0.785
Bartlett's Test (p-value)	1.64368E-87	4.90212E-87	3.42687E-92

From the Table 2, Pre-Peak-Post-COVID results of PCFA, the KMO test shows the adequacy of the factor analysis, and Bartlett's test of sphericity is highly significant. Both the tests confirm the adequacy of the PCFA and hence the extraction of common factors. Using varimax rotation, which maximizes factor variance, it is possible to obtain the factors based on eigenvalue greater than 1, that is, the selection of three factors each in pre and peak periods and two factors in post-period of COVID.

After the formation of the factors, the HII scores are calculated and standardized in order to form the index.

TABLE 3. - *HII scores of States/UT Pre-Peak-Post COVID period*

REGION	STATE/UT	Pre_HII	Peak_HII	Post_HII
East	Bihar	0.3100	0.4741	0.4049
East	Jharkhand	0.1758	0.1774	0.1781
East	Sikkim	0.0101	0.0118	0.0121
East	Odisha	0.3961	0.4066	0.3968
East	Manipur	0.0499	0.0529	0.0486
East	Assam	0.4501	0.4389	0.4049
East	Nagaland	0.0374	0.0353	0.0324
East	Mizoram	0.0208	0.0176	0.0162
East	Tripura	0.0600	0.0576	0.0567
East	Arunachal Pradesh	0.0410	0.0394	0.0405
East	Meghalaya	0.0612	0.0582	0.0607
East	Chhattisgarh	0.3189	0.3390	0.3117
East	West Bengal	0.5243	0.3226	0.3117
West	Gujarat	0.5867	0.6034	0.5628
West	Maharashtra	0.8789	0.8455	0.7895
West	Madhya Pradesh	0.4103	0.4395	0.4737

(follows)

West	Goa	0.0214	0.0170	0.0162
West	DandD Dadra and Nagar Haveli	0.0042	0.0035	0.0041
North	Ladakh	0.0053	0.0223	0.0121
North	Jammu and Kashmir	0.3688	0.3361	0.3239
North	Punjab	0.2245	0.2356	0.2308
North	Haryana	0.1597	0.1557	0.1498
North	Himachal Pradesh	0.1069	0.1093	0.1012
North	Rajasthan	0.7803	0.8519	0.8907
North	Uttarakhand	0.0689	0.0758	0.1053
North	Delhi	0.0038	0.0022	0.0036
North	Uttar Pradesh	0.8963	0.8620	0.8918
South	Tamil Nadu	0.7518	0.7814	0.7126
South	Kerala	0.2987	0.3296	0.3239
South	Karnataka	0.8682	0.7861	0.8178
South	Andhra Pradesh	0.4572	0.4871	0.4777
South	Telangana	0.3147	0.3120	0.2065
South	Puducherry	0.0143	0.0153	0.0081
South	Andaman and Nicobar Islands	0.0083	0.0094	0.0081
South	Lakshadweep	0.0047	0.0023	0.0040

From Table 3, it is clear that the highest HII for Pre-Peak-Post-COVID is found in the state of Uttar Pradesh (0.8962, 0.8519 and 0.8917), followed by the state of Maharashtra (0.8788), Rajasthan (0.8519 and 0.8906), and the lowest indices were registered in Delhi (0.00378, 0.00221 and 0.00364) respectively.

To check how the indices are spread across the country, the Hierarchical clustering of Ward's linkage with the application of Euclidean Distance is calculated for HII. For the years of 2019-20, 2020-21, and 2021-22 QGIS software is used to plot the HII into 7, 5, and 6 groups respectively, that are spatially spread across the States/UTs of India.

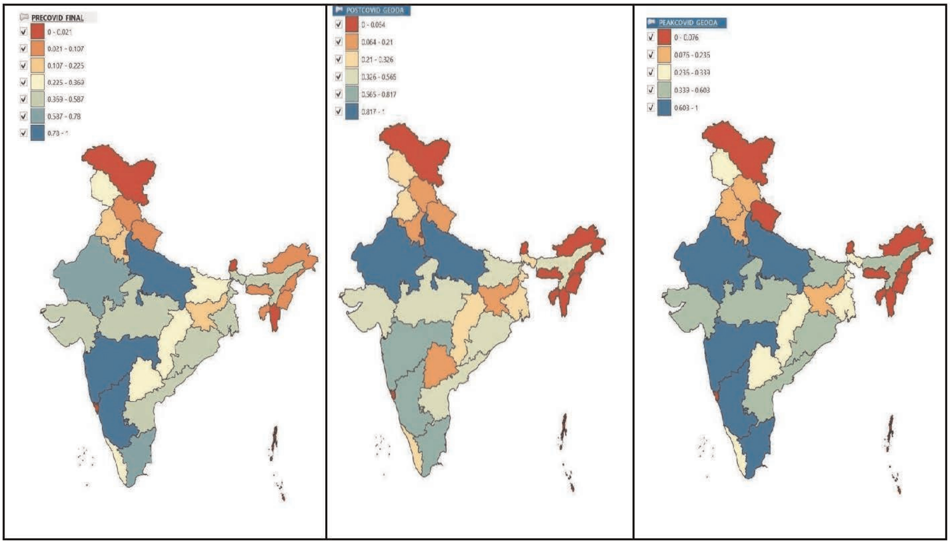

FIGURE 1. - *Spatial distribution of indices by groups - Pre, Peak, and Post COVID*

Figure 1 shows the groups of red shade indicating the lowest indices and the Blue shade indicating the highest indices. To check for the existence of patterns of spatial dependence and to detect the spatial autocorrelation, univariate Local Moran's I using distance weights of KNN algorithm is calculated. Based on the results, the null hypothesis of spatial randomness is rejected, confirming spatial autocorrelation in the dataset of HII scores. After confirming spatial dependence, it is possible to analyze local patterns of spatial association in the data using Local Moran's I. Cluster Map, Significance Map and Moran's I Scatter Plots are formed.

The results show the formation of high-high and low-low clusters in the Cluster Map. Low-low clusters mean that the states with a low HII are surrounded by other states with low HII. High-High clusters means the states with better healthcare infrastructure being surrounded by other states with similarly high levels of infrastructure. Figure 2 displays these cluster patterns across all three periods.

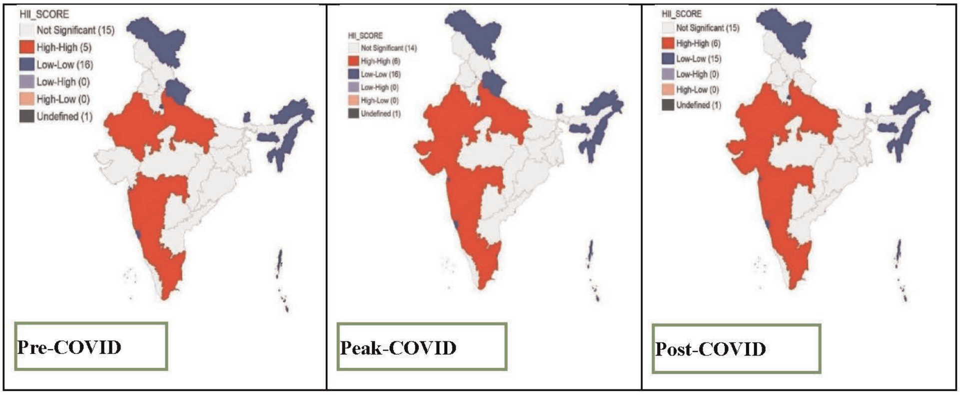

FIGURE 2. - *Cluster Maps of Pre-Peak-Post COVID*

Now for the Significance Map, the Clusters thereby formed on Cluster Map exhibit how much they are statistically significant through p values of 5%, 10% and 1% level of significance. The areas covered in the Clusters such as High-High or Low-Low clusters either fall on one of the significant levels of area covered in the Significance map plotted in Figure 3. This is to be understood that the areas received on the Cluster map - Cluster formation are proved statistically significant by the Significance Map plotted.

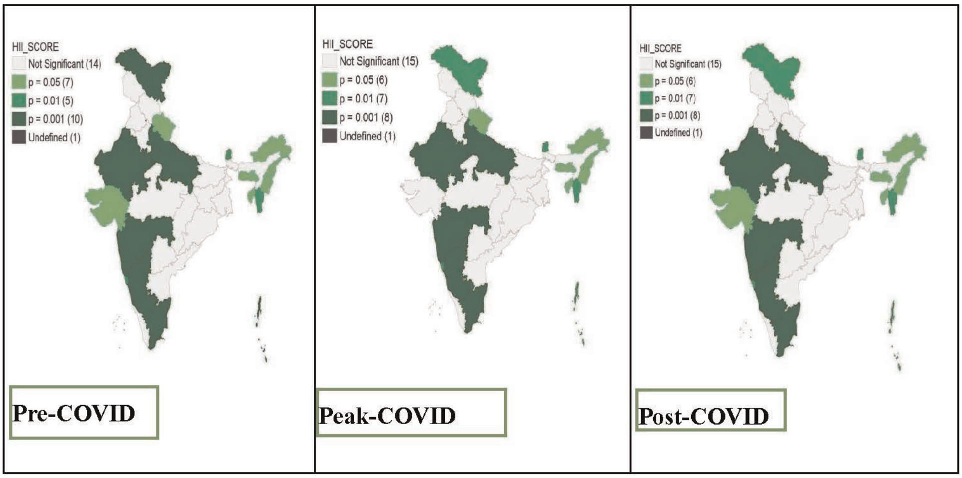

FIGURE 3. - *Significance Map of Pre, Peak and Post COVID*

Moran's I scatterplot shown in Figure 4 shares four quadrants where the first and third quadrants contribute to High-High and Low-Low clusters and the second, and fourth quadrants contribute to High-Low, and Low-High clusters respectively. Here from the scatterplot, we could see clusters being spread around first and third clusters thereby indicating the presence of High-High and Low-Low clusters respectively. The X-axis represents the HII score of the observations and Y-axis represents the spatial lagged values of the HII score. Moran's I statistic values are 0.949, 0.937, and 0.955 which are closer to 1, indicating strong positive spatial autocorrelation.

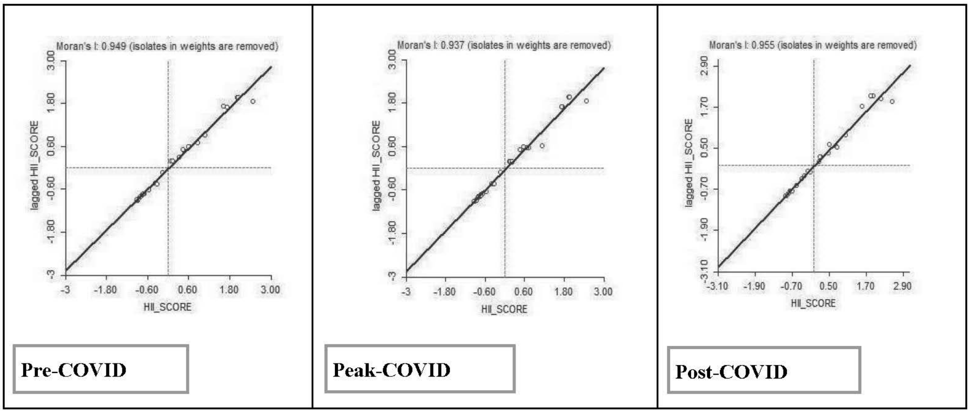

FIGURE 4. - *Moran's I Scatter Plots of Pre, Peak, and Post-COVID*

7. CONCLUSIONS

The COVID-19 outbreak in India presented significant challenges, particularly for the rural healthcare infrastructure, due to factors such as population density, inadequate healthcare facilities, and socio-economic disparities. Despite these obstacles, India's commitment to providing essential healthcare services to its rural population showcased remarkable resilience and determination.

The comprehensive analysis of healthcare infrastructure performance across various states and UTs during the three phases of the COVID-19 pandemic revealed notable insights. Using the PCFA, the study examined key variables such as staffing and primary infrastructure, highlighting the stability and adaptability of healthcare systems over time. The findings identified high- performing states like Uttar Pradesh, Rajasthan, Maharashtra, Tamil Nadu, and Karnataka, which demonstrated sustained excellence in healthcare provision. Conversely, regions like Lakshadweep, Delhi, Andaman and Nicobar Islands, Dadra and Nagar Haveli, and Goa showed lower performance levels, underscoring the need for targeted intervention and support.

The study emphasized that effective healthcare delivery transcends population metrics, relying instead on the equitable distribution, accessibility, and utilization of healthcare facilities. The success of high-performing states in meeting rural healthcare needs underscores the importance of equitable resource allocation and efficient service delivery mechanisms.

Government healthcare schemes in India have shown consistent improvement, particularly in selected states, reflecting a dedicated effort to enhance services, especially in rural areas; it includes developing infrastructure and launching many health insurance schemes. The study of HII helped the government in this regard and this study also threw lights on the need of such things. However, there remains a need for a more targeted approach at the State and UT levels to address specific healthcare challenges and Infrastructure developments. By concentrating on individual State's and UT's requirements, the government can ensure a better distribution of healthcare services and a stronger focus on rural populations.

In conclusion, the findings highlight the critical need for equitable distribution and effective utilization of healthcare resources to address the needs of all citizens. By advocating for targeted interventions and sustained support, the study aims to pave the way toward achieving universal health coverage and advancing the SDGs of the country. These efforts aspire to foster a healthier, more resilient nation where every individual has access to the healthcare services they need to thrive.

REFERENCES

Abdi H. (2003). Factor rotations in factor analyses. In M. Lewis-Beck, A. Bryman and T. Futing (Eds.), *Encyclopedia of Social Sciences Research Methods* (pp. 792-795). Thousand Oaks, CA: Sage.

Anand M. (2014). Health status and health care services in Uttar Pradesh and Bihar: A comparative study. *Indian Journal of Public Health*, **58**(3), 174-179.

Anselin L. (1995). Local indicators of spatial association (LISA). *Geographical Analysis*, **27**(2), 93-115.

Anselin L., Rey S.J. (2014). *Modern Spatial Econometrics in Practice*. Cambridge University Press, Cambridge.

Bezerra É.C.D., Santos P.S.D., Lisbinski F.C., Dias L.C. (2020). Spatial analysis of Brazil's COVID-19 response capacity: A proposal for a Healthcare Infrastructure Index. *Ciencia and Saude Coletiva*, **25**(12), 4957-4967.

Fávero L.P., Belfiore P. (2019). *Data Science for Business and Decision Making*. Academic Press, London.

Gatrell A.C, Bailey T.C. (1996). Interactive spatial data analysis. *Social Science and Medicine*, **42**(6), 843-855.

Gere A. (2023). Recommendations for validating hierarchical clustering in consumer sensory projects. *Current Research in Food Science*, **6**, 100522.

Goel M.M. (2011). *Economics of Human Resource Development in India*. VK Global Publications, New Delhi.

Mathur M. (2015). Spatial autocorrelation analysis in plant population: An overview. *Journal of Applied and Natural Science*, **7**(1), 501-513.

Paul P.K., Jana S.K., Maiti A. (2019). An analysis of the health status of the state of Assam, India. *Research Review International Journal of Multidisciplinary*, **4**(9), 1179-1188.

Rajeeb S. (2015). Block-level disparity in social development: A case study of Paschim Medinipur, West Bengal, India. *International Journal of Scientific Engineering and Research (IJSER)*, **3**(4), 92-95.

Rural Health Statistics (2018-2019). *Government of India, Ministry of Health and Family Welfare, Statistics Division.*

Trendafilov N., Gallo M. (2021). Principal component analysis (PCA). In *Multivariate Data Analysis on Matrix Manifolds: (with Manopt)* (pp. 89-139). Springer.

Trendafilov N., Gallo M. (2021). Factor analysis (FA). In *Multivariate Data Analysis on Matrix Manifolds: (with Manopt)* (pp. 141-186). Springer.

Tsai P.J., Lin M.L., Chu C.M., Perng C.H. (2009). Spatial autocorrelation analysis of health care hotspots in Taiwan in 2006. *BMC Public Health*, **9**, 464.