Paper submitted to International Population Conference 2025

Geographical Heterogeneity of Zero-Dose (Unvaccinated) Children in India: A Multilevel Geospatial Analysis of Cross-Sectional Survey Data

Mriganka Dolui*, Doctoral Fellow (PhD.), Department of Geography, Central University of Karnataka, 585367, India, email: mriganka.dolui@gmail.com; 21dpgeo04@cuk.ac.in

Sanjit Sarkar, Assistant Professor, Department of Geography, Central University of Karnataka, 585367, India, email: sanjitiips@gmail.com; sanjitsarkar@cuk.ac.in

Short Abstract

Routine immunization for children in India is a strategic priority to prevent child morbidity and mortality. Children who fail to receive the first dose of DTP (diphtheria, tetanus, and pertussis)-containing vaccine are known as zero-dose children due to limited healthcare access, sociodemographic factors and geographical spatiality. The present study aimed to decipher the spatial dependence, heterogeneity and determinants of zero-dose using geocoded data from 707 districts in India from the fifth round of the National Family Health Survey, 2019-21. Moran's I statistics and Gatis-Gi were used to identify spatial clustering and the degree of clustering (hotspot) of zero-dose, respectively. Multiple regression models, including ordinary least squares (OLS), spatial lag model (SLM), spatial error model (SEM), and geographically weighted regression (GWR), were employed to assess the geographic variation in determinants of children's zero dose. The overall prevalence of zero-dose was 6.67%, with a spatial dependence of 0.1753. GWR demonstrated the best model performance than OLS, SLM and SEM, revealing that the relationship between zero-dose and its determinants varied across districts in both magnitude and direction. The findings highlight the importance of spatial factors in understanding immunization gaps and can help policymakers in designing region-specific interventions to improve vaccination coverage across India.

Keywords: Zero dose, child mortality, Spatial analysis, geographically weighted regression, India

***Corresponding author and presenter:** Mriganka Dolui, E-mail: mriganka.dolui@gmail.com

Extended Abstract

Introduction

Children's vaccination is one of the most effective public health measures, preventing child morbidity and mortality from diseases like diphtheria, tetanus, pertussis (DTP), polio, measles, and Hepatitis-B (Hogan & Gupta, 2023; Singh, 2020). The World Health Organization's (WHO) extended Immunization Program has saved over 154 million lives and reduced infant mortality by 40% (WHO, 2019). The DTP vaccine, first developed in 1948, merged three antigens into one, reducing the number of required shots and increasing coverage. Currently, different combinations of DTP-containing vaccinations often include Hepatitis B, Haemophilus influenzae type B, and inactivated polio vaccine (UNICEF, 2023a). Despite these advancements, significant disparities in vaccination persist, particularly in lower-middle-income countries (LMICs). In 2022, 14.3 million children missed their first DTP dose, classified as "zero-dose" children, with another 6.2 million only partially immunized, indicating limited access to immunization and healthcare services (WHO, 2023). Moreover, another 6.2 million infants received only partial vaccination. Around 60% of zero-dose children reside in 10 countries, including India, Nigeria, Pakistan, Angola, Brazil, the Democratic Republic of the Congo, Ethiopia, Indonesia, Mozambique, and the Philippines (MoHFW & GOI, 2024; WHO, 2023)

Zero-dose children who have not received the first DTP-containing vaccine reflect limited healthcare access, while under-vaccinated children, those who missed DTP3, indicate gaps in immunization programs (UNICEF, 2023b). In India, DTP is administered as a pentavalent vaccine at 6, 10, and 14 weeks, with a booster at 16-24 months (MoHFW & GOI, 2017). Two-thirds of zero-dose children live in extreme poverty, highlighting the need for targeted interventions to meet Sustainable Development Goals (SDGs) like poverty reduction (SDG-1), better health outcomes (SDG-3), and diminished inequalities (SDG-10) (Hogan & Gupta, 2023). Global estimates revealed that DTP-1 coverage was at 94% in 2019, but the COVID-19 pandemic raised the number of unvaccinated children from 1.4 to 2.9 million by 2020 (UNICEF, 2023b). Therefore, this study reveals the geographical and sociodemographic correlates in the children zero-dose across districts in India.

Methods

Data

This study utilized the most recent round of the National Family Health Survey (NFHS-5; 2019-2021) data by the Health Ministry, Government of India (ICF & IIPS, 2021). The present study restricted to 41,132 children aged 12-23 months after excluding all the missing values in different socioeconomic characteristics. The 707 districts are considered as a unit of analysis. The district-level geocoded data were retrieved from the Demographic and Health Surveys (DHS) spatial repository https://spatialdata.dhsprogram.com.

Outcome variables

Children aged 12-23 months were considered to have received all basic vaccinations. The outcome variable was children's "zero doses", defined as all alive children from the 12-23 months age group who did not receive the first dose DPT (DTP1) or else considered as vaccinated (UNICEF, 2023b).

Explanatory variables

This study used a set of explanatory variables to explore the determinants of zero-dose children aged 12-23 months. Based on previous literature, we have considered the children's age (11-15, 16-19, 20-23months), sex of the child (male, female), mother's age (15-22, 23-30, 31-38, 39-49 years), place of delivery (home, public, private), Antenatal care (ANC) visits (no ANC visits, <4 visits, >4 visits), caste (Scheduled

Castes/SC, Scheduled Tribes/ST, Other Backward Classes/OBC, general), religion (Hindu, Muslim, and others religion), place of residence (urban, rural), mother's education (No education, primary, secondary, Higher), wealth (poorest, poorer, middle, higher), distance to the health centre (no problem, big problem), and media exposure (no, yes) as explanatory variables in the study (Fekadu et al., 2024; Kumar et al., 2016; Srivastava et al., 2020).

Statistical analysis

To assess the spatial distribution of children's zero-dose, district-level prevalence map was introduced across districts in India. Further, Moran's I statistics and Gatis-Gi were used to identify spatial autocorrelation (cluster) and quantify the degree of clustering (hotspot) of zero-dose, respectively. Moran's I is based on Waldo Tobler's first law of geography: "Everything is related to everything else, but close things are more related than distant things" (Tobler, 1970). Accordingly, zero dose prevalence should be similar among neighbouring districts to non-neighbours. Moran's I values range from -1 to 1: positive values show clustering, zero indicates randomness and negative values suggest dissimilar neighbouring values. Similarly, the Getis-Ord Gi* statistic measures the degree of clustering, with higher or lower z-scores indicating strong clustering of high or low values and classifies them into hotspots (high values next to high) and "coldspots" (low values next to low) (Anselin et al., 2007; Esri, n.d.).

Furthermore, multiple regression models, including ordinary least squares (OLS), spatial lag model (SLM), spatial error model (SEM), and geographically weighted regression (GWR), were employed to assess the geographical variation in determinants of children's zero dose. The OLS model was performed to identify predictors of the zero-dose with the coefficients that should be statistically significant and have either a positive or negative sign. Whereas SLM shows the dependency between the outcome and explanatory variables, and SEM postulates spatial dependence explains the error term of OLS (Ward & Gleditsch, 2018).

Moreover, global models like OLS, SLM, and SEM provide a single equation and average estimates for relationships across an entire dataset, potentially masking spatial variations. In contrast, local models like GWR offer spatial fluctuating parameter estimates, allowing relationships between variables to differ across locations. This flexibility helps better capture the heterogeneity of associations, addressing limitations in global models that assume constant relationships across space (Fotheringham et al., 2017). GWR thus enables location-specific insights into variable relationships. Regarding model optimization to select the best model, Akaike information criterion (AICc) was used, and the bandwidth with the lowest AICc was selected for the model. Additionally, the variance inflation factors (VIFs) were calculated to detect multicollinearity, and no such collinearity exists in the involved model where all the VIF values are less than 10. The statistical and spatial analysis was conducted on Stata (Stata-17) and ArcMap software (ArcMap 10.8.1) (ESRI, 2011; StataCorp, 2021).

Results

Prevalence of zero doses children across socioeconomic background

The results revealed that the overall prevalence of zero-dose children was 6.6%. By and large, the prevalence of the zero doses was higher among children who were born at home (12.96%), who had no ANC visits (12.03%) and children from ST (7.82%), mothers with no education (10.43%), belonged to the poorest households (8.85%), and no media exposure among mothers (9.36%) (Figure 1).

Spatial variation and Spatial autocorrelation of zero doses

Figure 2a shows the spatial variation of children's zero dose across several districts in India. The highest zero-dose prevalence ranges from 19.61%-41.82% (17 out of 707 districts) . The spatial distribution of zero-dose children showed a significant spatial variation across districts with Moran's I value of 0.1753 (p-value <0.001) (Figure 2b). The associated Z-score was 25.77, elucidated a clustered pattern in zero-dose, indicating children's zero-dose are significant and are positive spatial autocorrelated. Figure 2c represents the hotspot Gatis-Gi map depicting the hotspot-coldspot spatial clustering revealed the higher hotspot of children's zero doses observed in northeastern and central districts in India.

Determinants of children's zero-dose

Global regression model: OLS, SLM and SEM

The OLS regression model identifies the predictors of children with zero doses, indicating that the OLS model can explain 32.2% (R-squared: 0.322) of the variance in zero doses of the children (Table 1). Whereas the spatial regression models (SLM and SEM) explicitly outperform their non-spatial counterparts, as evidenced by higher R-squared and lower AIC values. The SEM outperformed the SLM. As such, the following interpretation is based on SEM because it seems more reasonable than the OLS and SLM. The SEM results explicitly revealed that no ANC visits (0.1705; p-value <0.05), less than 4 ANC visits (0.0321; p-value <0.05), and mothers with no media exposure (0.0966; p-value <0.05) have significant and positive predicts children's zero doses in India. Whereas children whose place of delivery was in public institutions (-0.0640; p-value <0.05), belonging to the Hindu religion (-0.0223; p-value <0.05), and residing in rural areas (-0.0222; p-value <0.05) show a significant and negative effect on zero-dose children (Table 1).

Local regression analysis: GWR model

Table 2 shows the GWR modelling results for zero-dose children in India. The results suggest that the GWR model has better explanatory power and can explain more variations in the predicted variable than the linear regression model due to lower AIC value (4179.320) and higher R-squared value (0.358) (Table 3). Furthermore, Figure 3 represents the spatial distribution of the coefficients of the selected predictors of zero doses from the GWR model, suggesting that the magnitude of the coefficients varies spatially. For instance, public deliveries predicted higher zero-dose rates in the northeast (Figure 3a), while lack of ANC visits (no ANC visits and <4 ANC visits) was significant in the northeast and southwest (Figure 3b, 3c). Zero-dose rates were also higher among children from Hindu families in the west (Figure 3d), rural areas in the south (Figure 3e), and mothers with no media exposure in the north and some southern regions (Figure 3f).

Conclusion

The study underscores the importance of focusing on vulnerable regions and geographic hotspots to improve immunization coverage. On the one hand, spatially fixed policies cannot be proposed for zero-dose children to eliminate it. Policy defragmentation is a great concern in rural areas and government or public hospitals regarding vaccination. On the other hand, ANC visits of mothers are another concern, where awareness and grassroots health workers need to be more active at the local level to cater to each child for immunization. Increasing media outreach and digital campaigning are needed in a vibrant manner for vaccination awareness to reduce the zero-dose burden and achieve more equitable health outcomes for children in India.

Tables:

Table 1: OLS and spatial regression models (SLM and SEM) show estimated coefficients for the children's zero-dose in India, 2019-21

Variables	Zero Dose					
	OLS		SLM		SEM	
	Coefficient	p-value	Coefficient	p-value	Coefficient	p-value
Child age- 16-19 months	0.0084	0.7666	0.0084	0.7480	0.0008	0.9769

Child age- 20-23 months	-0.0366	0.2448	-0.0313	0.2218	-0.0291	0.2511
Child- male	0.0051	0.8373	0.0013	0.9563	0.0010	0.9640
Mother's age- 23-30 years	-0.0043	0.8436	-0.0037	0.8651	-0.0032	0.8888
Mother's age- 31-38 years	0.0265	0.3119	0.0270	0.3043	0.0348	0.2117
Mother's age- 39-49 years	0.0653	0.4533	0.0623	0.4568	0.0931	0.2610
Place of delivery-public	-0.0648	0.0244	-0.0564	0.0095	-0.0640	0.0055
Place of delivery-private	-0.0358	0.2771	-0.0303	0.2462	-0.0332	0.2346
No ANC visits	0.1574	0.0000	0.1534	0.0000	0.1705	0.0000
<4 ANC visits	0.0363	0.0112	0.0258	0.0566	0.0321	0.0259
Caste- SC	-0.0279	0.1429	-0.0257	0.1971	-0.0288	0.1699
Caste- ST	0.0130	0.3519	0.0106	0.4454	0.0039	0.7927
Caste- OBC	0.0022	0.8334	0.0016	0.8932	-0.0019	0.8900
Religion- Hindu	-0.0235	0.0428	-0.0169	0.1580	-0.0223	0.0952
Religion- Muslim	-0.0151	0.3283	-0.0095	0.5405	-0.0132	0.4478
Residence rural	-0.0217	0.0487	-0.0176	0.1141	-0.0222	0.0627
Mother education- No	-0.0235	0.3264	-0.0201	0.4094	-0.0212	0.4361
Mother education- primary	-0.0021	0.9482	-0.0101	0.7438	-0.0175	0.6013
Mother education- secondary	0.0264	0.1477	0.0230	0.2637	0.0226	0.3209
Wealth Index- poorest	-0.0227	0.2654	-0.0218	0.2855	-0.0118	0.5955
Wealth Index- poorer	0.0176	0.3718	0.0100	0.6183	0.0149	0.4866
Wealth Index- middle	0.0316	0.1686	0.0300	0.2010	0.0369	0.1302
Distance to a health facility- big problem	-0.0154	0.4189	-0.0116	0.4477	-0.0042	0.7958
No media exposure	0.1003	0.0000	0.0915	0.0000	0.0966	0.0000
Constant	9.5190	0.0347	7.5418	0.0711	9.4903	0.0300
Lag Coefficient (Rho/Lambda)			0.1996	0.0001	0.2765	0.0000
Spatial model performance statistics						
AIC	4210.9000		4197.630		4187.240	
R-squared	0.3226		0.342		0.355	
Log-likelihood	-2080.4500		-2072.810		-2068.619	

Table 2: Geographically weighted regression results in zero dose across demographic and socioeconomic characteristics in India, 2019-21

Variables	GWR coefficient			
	Min	**Median**	**Max**	**Standard deviation**
Place of delivery public	-0.072957	-0.028459	-0.001450	0.018846
No ANC visits	0.086792	0.484639	0.134798	0.031020
<4 ANC visits	-0.007961	0.045474	0.098025	0.016441
Religion Hindu	-0.090771	-0.028220	-0.000459	0.022287
Residence rural	-0.026962	-0.019210	0.061504	0.010960
No media exposure	-0.051999	0.060074	0.144092	0.029643
GWR performance statistics				
AICc	4179.320			
R-squared	0.358			
Adjusted R-squared	0.326			
Sigma	4.580			

Table 3: Comparison of OLS and GWR results in zero doses in India, 2019-21

Model comparison statistics	OLS Model	GWR Model
AICc	4214.965	4179.320
R-squared	0.322	0.358
Adjusted R-squared	0.299	0.326
Moran's I	0.175	0.379

Figures:

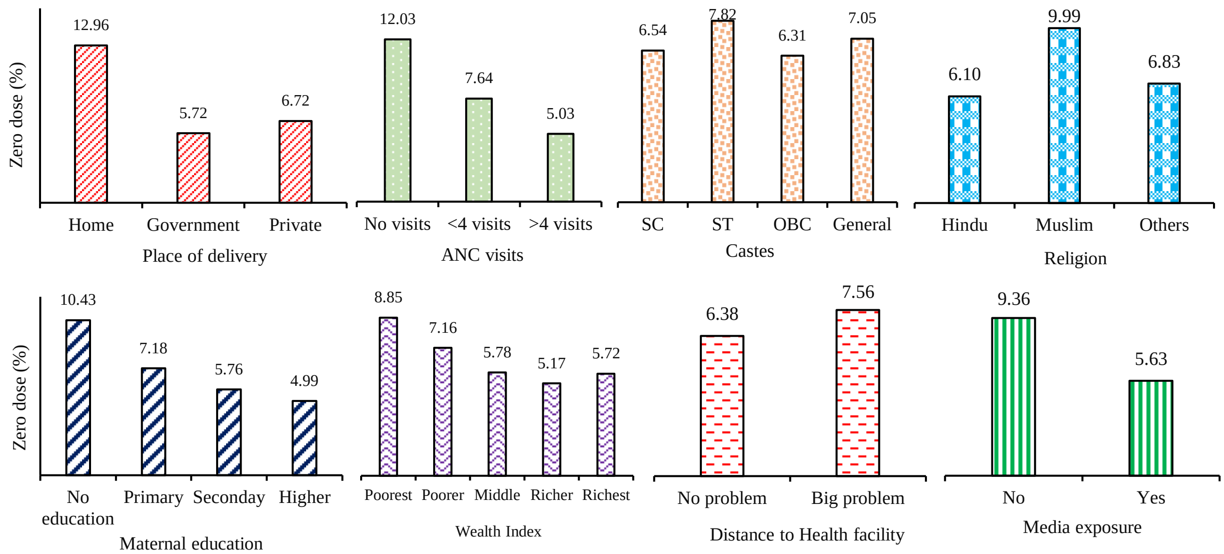

Figure 1: Prevalence of zero-dose children across different demographic and socioeconomic characteristics in India

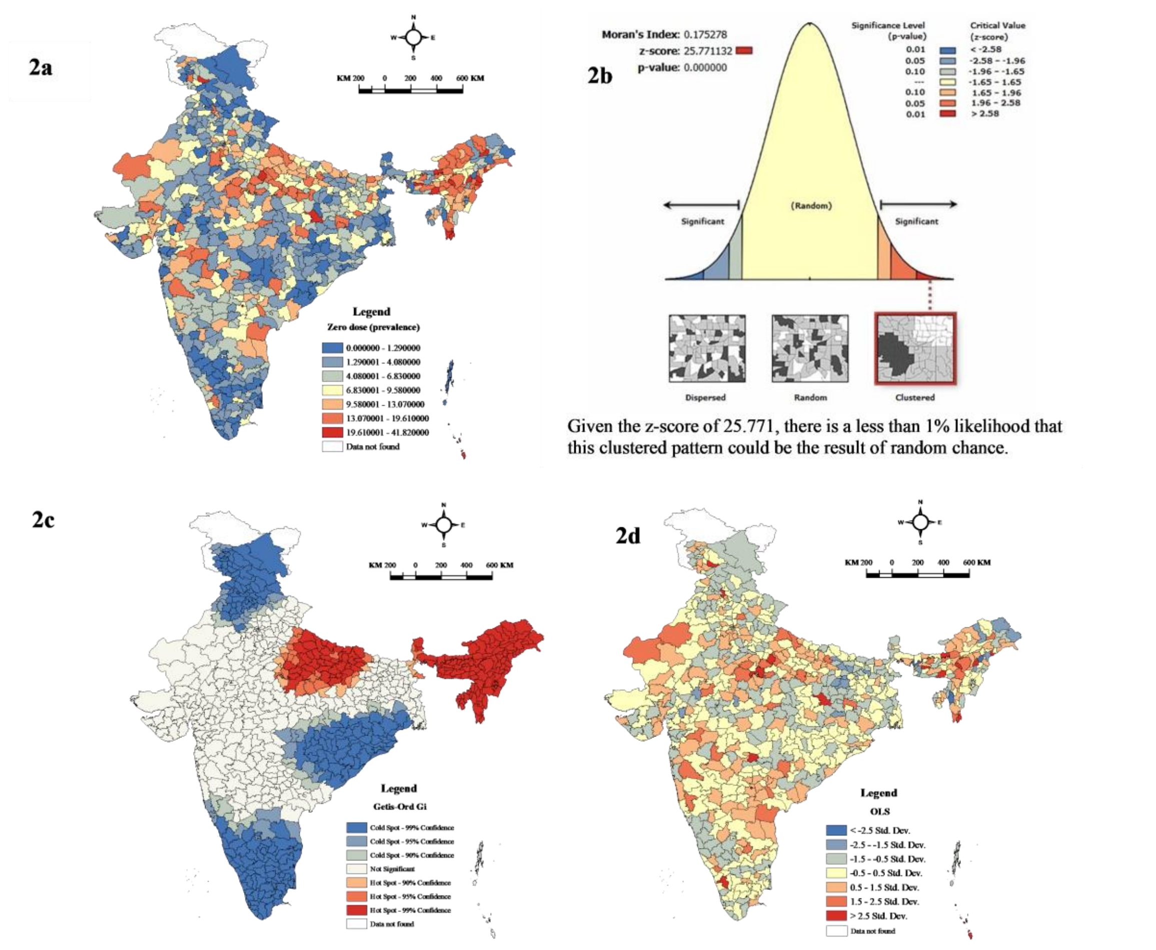

Figure 2: a) District level prevalence of children's zero dose, b) Moran's I, c) Hotspot Gatis-Gi map showing spatial clustering of children's zero doses across districts in India, and d) OLS coefficient

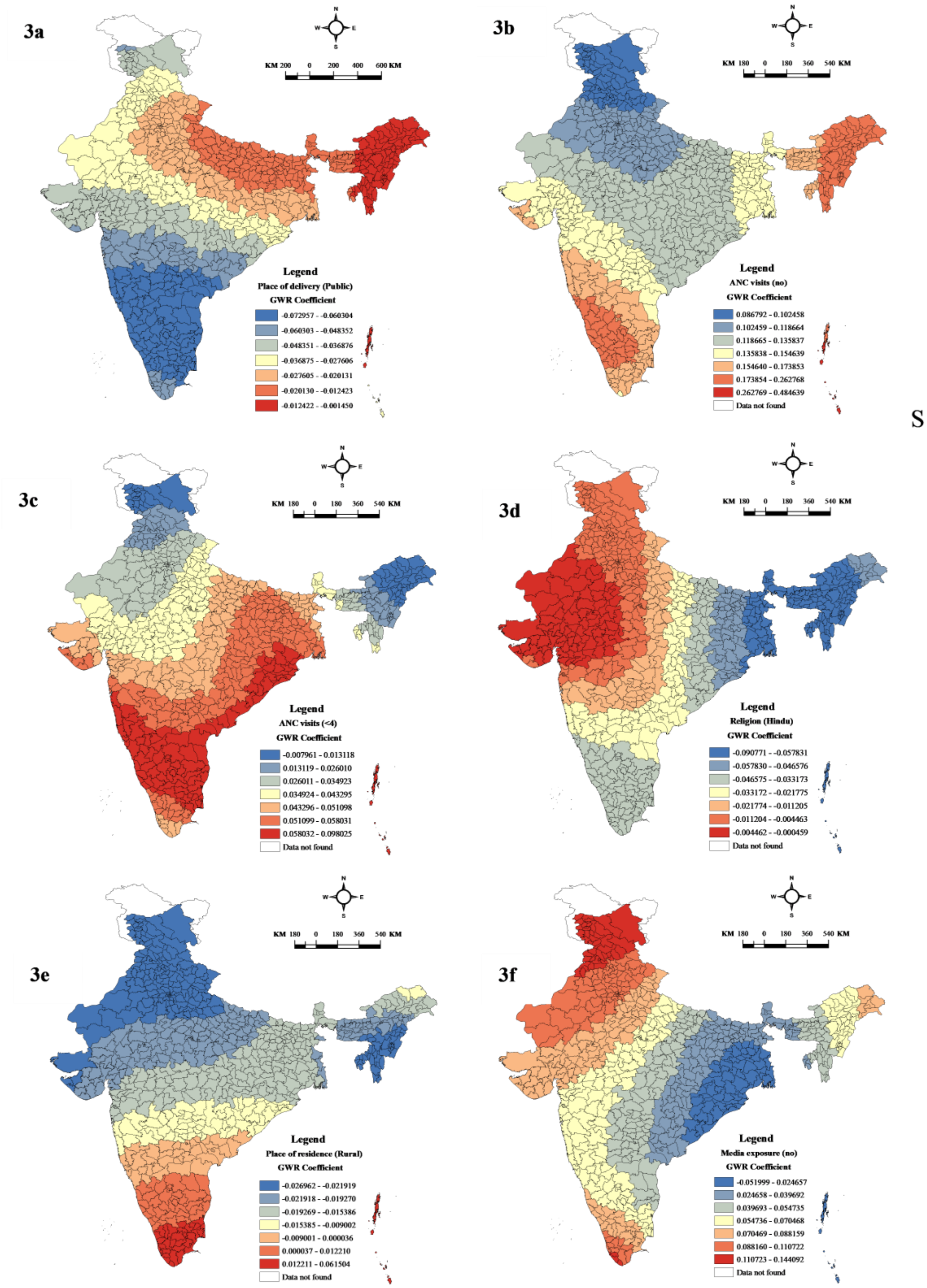

Figure 3: Spatial distribution of the GWR coefficient of selected predictors on zero dose among children in India

Note: The shapefile was downloaded from Spatial Data Repository (SDR) – Boundaries, The Demographic and Health Surveys Program (DHS), ICF International, funded by the United States Agency for International Development (USAID). Accessible via the website https://spatialdata.dhsprogram.com [Retrieved on April 01, 2024]. The SDR provides geographically linked health and demographic data from the DHS Program and the U.S. Census Bureau for mapping in a geographic information system (GIS).

References:

Anselin, L., Sridharan, S., & Gholston, S. (2007). Using exploratory spatial data analysis to leverage social indicator databases: The discovery of interesting patterns. *Social Indicators Research*, *82*(2), 287–309. https://doi.org/10.1007/S11205-006-9034-X/METRICS

Esri. (n.d.). *High/Low Clustering (Getis-Ord General G) (Spatial Statistics)*. Retrieved September 1, 2024, from https://pro.arcgis.com/en/pro-app/latest/tool-reference/spatial-statistics/high-low-clustering.htm

ESRI. (2011). *ArcGIS Desktop: Release 10. Redlands, CA: Environmental Systems Research Institute*.

Fekadu, H., Mekonnen, W., Adugna, A., Kloos, H., & Hailemariam, D. (2024). Trends of inequality in DPT3 immunization services utilization in Ethiopia and its determinant factors: Evidence from Ethiopian demographic and health surveys, 2000-2019. *PloS One*, *19*(1). https://doi.org/10.1371/JOURNAL.PONE.0293337

Fotheringham, A. S., Yang, W., & Kang, W. (2017). Multiscale Geographically Weighted Regression (MGWR). *Annals of the American Association of Geographers*, *107*(6), 1247–1265. https://doi.org/10.1080/24694452.2017.1352480

Hogan, D., & Gupta, A. (2023). Why Reaching Zero-Dose Children Holds the Key to Achieving the Sustainable Development Goals. *Vaccines*, *11*(4). https://doi.org/10.3390/vaccines11040781

ICF, & IIPS. (2021). *NATIONAL FAMILY HEALTH SURVEY (NFHS-5) 2019-21 INDIA*. http://www.rchiips.org/nfhs

Kumar, C., Singh, P. K., Singh, L., & Rai, R. K. (2016). Socioeconomic disparities in coverage of full immunisation among children of adolescent mothers in India, 1990–2006: a repeated cross-sectional analysis. *BMJ Open*, *6*(8), e009768. https://doi.org/10.1136/BMJOPEN-2015-009768

MoHFW, & GOI. (2017). *Immunization Handbook for Medical Officers 2017*.

MoHFW, & GOI. (2024). *Guidance Document on Strategic Approach for Reaching Zero Dose Children in India. India's Zero-Dose Implementation Plan*.

Singh, V. (2020). The Effect of COVID-19 Pandemic on Immunization Services in India - Possible Challenges and way forward. *Epidemiology International (E-ISSN: 2455-7048)*, *5*(2), 53–60. https://doi.org/10.24321/2455.7048.202018

Srivastava, S., Fledderjohann, J., & Upadhyay, A. K. (2020). Explaining socioeconomic inequalities in immunisation coverage in India: New insights from the fourth National Family Health Survey (2015-16). *BMC Pediatrics*, *20*(1), 1–12. https://doi.org/10.1186/S12887-020-02196-5/TABLES/6

StataCorp. (2021). *Stata: Release 17. Statistical Software. College Station, TX: StataCorp LLC*.

Tobler, W. R. (1970). A Computer Movie Simulating Urban Growth in the Detroit Region. *Economic Geography*, *46*, 234. https://doi.org/10.2307/143141

UNICEF. (2023a). *Diphtheria Tetanus and Pertussis Containing Vaccines: Market and Supply Update UNICEF Supply Division*.

UNICEF. (2023b). *The State of the World's Children 2023: For every child, vaccination*. https://www.unicef.org/reports/state-worlds-children-2023

Ward, M., & Gleditsch, K. (2018). *Spatial regression models*. https://books.google.com/books?hl=en&lr=&id=iyZaDwAAQBAJ&oi=fnd&pg=PT13&ots=0ob8WlT1YB&sig=skjwhKSg41VSUCmc1RO1Z1a0yRs

WHO. (2019, December 5). *Immunization*. https://www.who.int/news-room/facts-in-pictures/detail/immunization

WHO. (2023, July 18). *Immunization coverage*. https://www.who.int/en/news-room/fact-sheets/detail/immunization-coverage