Hotspot Clustering and Geographically Weighted Regression Analysis of Domestic Violence in Northeast India

Sharatchandra Haobijam[1]*, Kh. Jitenkumar Singh[2], Meena Hijam[3], Ksh. Anand Singh[1], Haobijam Nirendrakumar Singh[3], Nandeibam Alfred Rozer[3], and Arun Naorem[3]

[1] Department of Statistics, Manipur University, Manipur, India
[2] Indian Council of Medical Research, New Delhi, India
[3] Community Medicine Department, Jawaharlal Nehru Institute of Medical Sciences, Manipur, India
* Sharatchandra Haobijam, corresponding author. Email: haobijamsharatchandra@gmail.com
Submitted: 7 December 2023. Accepted: 10 August 2024. Published: 13 December 2024
Volume 33, 2025. pp. 742–761. http://doi.org/10.25133/JPSSv332025.040

Abstract

Domestic violence (DV) is now widely recognized as a severe public health problem owing to its health consequences. India has high prevalence rates of physical, sexual, and emotional violence against spouses (28%, 14%, and 6%, respectively). The study uses data from the National Family Health Survey (NFHS-5) to analyze the spatial distribution of different forms of DV in Northeast India. Bivariate analysis, ordinary least squares (OLS), and geographically weighted regression (GWR) were employed for data analysis. Domestic violence in Northeast India stands at 31.3%, with Manipur at 41.5%, followed by Assam and Arunachal Pradesh. Hailakandi in Assam (64.7%) and Bishnupur in Manipur (59.9%) have the highest rates. The local R^2 values for domestic violence were notably higher in the southern and eastern regions of the northeast States. Specifically, in the southeastern districts of Nagaland, these values ranged between 0.65 and 0.70. Regional disparities were evident in the prevalence of physical, emotional, and sexual violence, with Manipur, Assam, and specific districts in Arunachal Pradesh and Nagaland frequently highlighted as hotspots. The results highlight the necessity of region-specific strategies and focused interventions to effectively address and prevent DV throughout the Northeast. Prioritizing the mitigation of significant risk factors for DV in hotspot regions should be the government's top priority.

Keywords

Geographically weighted regression; hotspot clustering; NFHS-5; ordinary least squares

Background

Domestic violence (DV) is a societal problem that transcends boundaries of race, gender, culture, age, educational level, and socioeconomic status. Pregnant women are particularly vulnerable to DV (Jamshidimanesh et al., 2013). Violence against women is a universal phenomenon and has become a substantial public health concern; an estimated 30% of women experience intimate partner violence (IPV) globally (World Health Organization, 2021). Intimate partner violence is commonly categorized into three categories: Physical Violence (PV), Psychological or Emotional Violence (EV), and Sexual Violence (SV) (World Health Organization, 2005). It is a severe violation of women's human rights and is associated with poor physical and mental health outcomes (Devries et al., 2011; Potter et al., 2021). Further, women who are victims of IPV reported experiencing adverse reproductive health outcomes such as unintended pregnancy and abortion, especially in low-and middle-income countries (Pallitto et al., 2013).

According to the World Health Organization (WHO), the prevalence of IPV in any form varies across regions: 35% in Southern Asia, 33% in Sub-Saharan Africa, 25% in Northern America, and 23% in Northern Europe region (Shidhaye & Giri, 2014). Despite the varying prevalence rates, the risk factors experienced by women were similar across the 15 study sites of the WHO multi-country study on women's health and domestic violence: alcohol abuse by partner or self, women cohabitating with a partner without formal marriage, younger age of women, women supportive of wife beating, having sexual partners out of marriage, the experience of childhood abuse and domestic violence (Abramsky et al., 2011). The ecological model emphasizes that violence is created by a complex interaction of risk factors that operates at the family and individual level, along with broader community, cultural, and social variables. Low educational attainment, unemployment, and alcohol consumption are examples of factors that contribute to violence at the individual and family levels. On the other hand, violence is more likely to occur in communities that have patriarchal societies, high population densities, cultural norms that normalize violence, and areas lacking institutional support (Dahlberg & Krug, 2006).

In the Indian scenario, 32% of ever-married women reported having experienced domestic violence in their lifetime by their spouses. Physical, sexual, and emotional violence are the most common forms of spousal violence in India, with prevalence rates of 28%, 14%, and 6% respectively. Among these victims, 84% reported the husband as a perpetrator of domestic violence (International Institute for Population Sciences [IIPS] & ICF, 2022). In a systematic review of domestic violence among Indian women, varied prevalence rates of all forms of DV were reported (Kalokhe et al., 2017). This inter-study variance in the prevalence rates highlights the need to standardize the tools used to study domestic violence; otherwise, drawing a valid conclusion becomes difficult. Further, it also becomes essential to understand the spatial variation and clustering of domestic violence as it will help target areas with hotspot clusters.

This is particularly significant in Northeast India, one of the least developed regions in terms of economy, connectivity, and institutional infrastructure (Basumatary & Panda, 2020; Lahiri, 2017). It also has a unique topography in terms of social and cultural settings. Since the area is a part of the Golden Triangle, drug abuse and trafficking are among the highest (Baruah & Baruah, 2021). Moreover, Northeast India faces significant challenges with immigration and unemployment (Singh & Suman, 2016). These circumstances exacerbated DV in northeastern India.

The region has both the highest and lowest rates of domestic violence in the country—Manipur reported 55%, and Sikkim reported 3.5% (IIPS & ICF, 2017). Despite this alarming prevalence, limited studies in this region use spatial techniques. For instance, Indian studies using GIS techniques have shown spatial variation or discrepancy regarding the prevalence and determinants of social and health-related behaviors, such as alcohol use (Roy et al., 2024b), tobacco use (Singh et al., 2021), quality of life (Roy et al., 2023), and living conditions (Roy et al., 2024a). However, there is a shortage of studies that have looked at domestic violence through GIS technology, particularly in the Northeast region. Researchers have highlighted the need for employing spatial crime analysis, crime mapping, and visualization (Roy & Chowdhury, 2023b), which will further serve as a vital element in reducing the prevalence of violence. Thus, the objectives of the present study are (i) to analyze the spatial distribution of different forms of domestic violence in Northeastern India and (ii) to explore the underlying factors contributing to the spatial variations in domestic violence across this region. Identifying these spatial patterns and determinants will help inform policymakers and stakeholders in designing targeted interventions to reduce domestic violence and improve the well-being of affected communities. Underlying factors such as women's age, education, working status, wealth index, and husband's alcohol were found to be predictors of domestic violence. Hence, these variables will be considered for the study.

The rationale of the study

In India, there are several forms and degrees of DV. Under Indian Penal Code 498-A, such violence has been declared illegal since 1983. However, until the Protection of Women from Domestic Violence Act 2005 was passed in 2006, it did not provide any protection to victims of domestic abuse (Gazette of India, 2005). India has various sociocultural contexts in different parts of the country. As a result, when implementing intervention measures, studies on domestic violence that are state- or region-specific will be more pertinent. The results may help government programs and policy designers create efficient intervention strategies (Haobijam & Singh, 2023). To achieve this, the current study looks for a few explanatory variables that substantially impact DV in the Northeast Indian states. According to the NFHS-4, the states in Northeast India that have reported the worst and best-case scenarios are Manipur, Sikkim, Mizoram, Assam, Meghalaya, Nagaland, and Tripura (Haobijam & Singh, 2022). The present study is based on Northeast India, which either serves as a model for other states or clears the path for future research on all other states.

Data

Data for this study were obtained from the fifth round of the National Family Health Survey (NFHS- 5, 2019–2021) coordinated by the International Institute for Population Sciences (IIPS) and ICF International, Mumbai, conducted with support from the Ministry of Health & Family Welfare (MoHFW), Government of India. The survey provides updated and reliable information on crucial population indicators such as reproductive health, fertility, maternal and child mortality, high-risk sexual behavior, nutritional status of women and children, family planning methods, immunization, non-communicable diseases, women's autonomy, and DV. This is a nationally representative probability sample of all women aged 15 to 49 (*N*

= 724,115), from which 464,830 were administered the gender-based violence module. In total, 30,456 clusters (also called primary sampling units) were selected to collect samples, of which fieldwork was done in 30,198 clusters. The 2011 Census enumeration served as the sampling frame for the selection of clusters. In the first stage, the clusters were selected using the probability proportional to size (PPS) method. In the second stage, a complete household mapping and the listing were done in the selected clusters, and 22 households were randomly picked up in each cluster from the household listing. The DV module interviewed only one eligible woman per household, randomly selected to answer questions in the DV section to comply with ethical requirements.

A detailed description of the sampling design and survey procedure is provided in the NFHS-5 national report (IIPS & ICF, 2022). In Northeast India, 103,433 women aged 15–49 were interviewed, of which only 11,246 were selected and interviewed for the DV module. Out of 11246 women, only 9582 women are ever married, on which the present analysis is carried out. The present study excluded never-married women in the age group 15–49. Northeast India comprises eight states: Anurachal Pradesh, 1,779 women in the age group 15–49 were selected and interviewed for the DV module, along with 3,394 women from Assam, 729 from Manipur, 1,153 from Meghalaya, 635 from Mizoram, 801 from Nagaland, 280 from Sikkim, and 811 from Tripura.

Dependent variable

The present study considered domestic violence as a dependent variable. It also provides information on three types of violence against women: physical, sexual, and emotional. The violence was measured by asking all ever-married women if their husbands ever committed the following to them:

Physical violence: Pushing, shaking, throwing something, slapping, punching or hitting by a harmful object, kicking or dragging, strangling or burning, threatening with a knife or gun or any weapon.

Emotional violence: Ever been humiliated by husband/partner, ever been threatened with harm by husband/partner, ever been insulted or made to feel inadequate by husband/partner.

Sexual violence: Ever been physically forced into unwanted sex by husband or partner, ever been forced into other unwanted sexual acts by husband or partner.

Independent variables

Seven independent variables were selected that are highly significant for DV in Northeast India: the age of women, women’s education, working status of women, marital status of women, wealth index, husband’s drinking habit, and religion (Haobijam & Singh, 2022).

Statistical analysis

Bivariate analysis was used to estimate the prevalence of DV in Northeast states. Further, we computed the proportion of women aged 35–49 years (DV is less prevalent among respondents aged 15–34 years compared to those aged 35 years and older (Haobijam & Singh, 2022), the proportion of women who follow the Christian religion (Muslims have a higher percentage of DV incidents, but the distribution of the Muslim population is not uniform across all states, as evidenced by states like Sikkim and Mizoram (Haobijam & Singh, 2022), the proportion of women based on their working status, the proportion of divorced or widowed or separated women, the proportion of women whose husbands drink alcohol, the proportion of illiterate women, and the proportion of women in the poorest category. We used STATA (Version 18; StataCorp), employing the *svy* command for complex survey design, R (version 4.3.2), and QGIS for data analysis. Further, we used global Moran's *I*, which indicates a dataset's overall spatial autocorrelation. The second measure is a Local Indicator of Spatial Auto-correlation (LISA) measure local Moran's *I*, which means the "presence or absence of significant spatial clusters or outliers for each location" in a dataset.

Global spatial autocorrelation, measured by Moran's *I*, captures the extent of overall clustering or quantifies the degree of spatial autocorrelation in a dataset across all the districts. LISA essentially measures the statistical association between the value in District *I* and the value of the nearby district. A positive LISA statistic identifies a spatial concentration of similar values. When the LISA statistic is negative, we have a spatial cluster of dissimilar values, such as an area with high outcome values surrounded by low-outcome values. The Hotspot Analysis tool computes the Getis-Ord G_i^* statistic for each feature in the dataset. The resultant Z score of the G_i^* statistic shows where the features spatially cluster with either high or low values. A high *z*-score and a low *p* value indicate a significant hotspot; a low negative z-score and a small *p* value indicate a significant cold spot. A z-score near 0 means no spatial clustering.

Ordinary least square and geographically weighted regression

We performed the ordinary least square (OLS) regression, also known as Global regression, to identify the significant predictors of the observed spatial pattern of domestic violence (physical, emotional, and sexual). Before running the OLS model, we checked for all required assumptions.

The regression equation can be expressed as:

$$y_i = \beta_0 + \sum \beta_k x_{ki} + \varepsilon_i \quad (1)$$

Where y_i is the dependent variable, β_k the coefficients, x_{ki} is the independent variable, and ε_i is the error term.

Geographically weighted regression (GWR), an extension of OLS regression that models relationships as they vary across space by evaluating where locally weighted regression coefficients deviate from global coefficients, was used to ensure the heterogeneity of coefficient across each cluster of Northeast region that examine how those relationships between outcome variable, i.e., DV (physical, emotional, and sexual) and explanatory variables vary spatially (Roy et al., 2024b). Unlike OLS, which fits a single linear equation for all the data in the study area, GWR creates an equation for each cluster. While the equation in OLS is calibrated using data from all clusters, GWR uses data from nearby clusters. Therefore, GWR coefficients take different values for each cluster—maps of the β-coefficient associated with each independent variable guide the targeted interventions. We applied the Jarque-Bera test ($p < .001$) to verify the residual normality assumption. As residuals are not spatially correlated, the Koenkar BP test ($p < .001$) was conducted to check if the model was undergone for GWR. In addition, multicollinearity was checked using variance inflation factor (VIF). The VIF<5 shows no significant multicollinearity.

This method generates a separate regression equation for each observation, which can be expressed as follows:

$$y_i = \beta_0(u_i,v_i)+\sum \beta_k(u_i,v_i)\, x_{ki} + \varepsilon_i \quad (2)$$

Where y_i is the dependent variable, β_k the coefficients, x_{ki} the independent variables, (u_i, v_i) the coordinate location of i and ε_i is the error term.

We fitted the following GWR model proportion of domestic violence (physical, emotional, sexual and severe violence) = $\beta_0(x_i, y_i)+\beta_1(x_i, y_i)X_1+ \beta_2(x_i, y_i)X_2+\beta_3(x_i, y_i)X_3+\beta_4(x_i, y_i)X_4+\beta_5(x_i, y_i)X_5+\beta_6(x_i, y_i)X_6+\beta_7(x_i, y_i)X_7$ where x_i and y_i are the spherical x-y coordinates and

X_1: Proportion of women aged 35–49 years
X_2: Proportion of women in the Christian religion
X_3: Proportion of women with working status
X_4: Proportion of divorced/widow/separated women
X_5: Proportion of women's husbands drink alcohol
X_6: Proportion of illiterate women
X_7: Proportion of women in the poorest

We used an adaptive bisquare spatial kernel, with bandwidth optimized using the Golden Section method and the corrected Akaike information criterion (AICc) as the optimization criterion.

Results

In Northeast India, the overall prevalence of domestic violence (DV) stands at 31.3%, with Manipur recording the highest rate at 41.5%, followed by Assam, Arunachal Pradesh, Tripura, Meghalaya, Sikkim, Mizoram, and Nagaland. Table 1 presents the prevalence of various forms of DV, including physical, emotional, and sexual violence, reported by ever-married women in Northeast India.

Table 1: Percentage Distribution of Different Forms of DV in Northeast India

State	Domestic Violence (DV)	Physical Violence (PV)	Emotional Violence (EV)	Sexual Violence (SV)	Sample (n)
Arunachal Pradesh	26.5	23.8	12.9	6.2	1,779
Assam	34	31.3	11.8	7	3,394
Manipur	41.5	38.5	11	4.9	729
Meghalaya	21.1	13.2	13.7	6.1	1,153
Mizoram	11.9	9.9	5.9	1.9	635
Nagaland	11	6.1	7.6	0.9	801
Sikkim	21	10.6	14.6	2.7	280
Tripura	23	19.3	11.3	6.1	811
Total	31.3	28	11.7	6.4	9,582

Note: The author calculated this using NFHS-5 data from the MoHFW, Government of India.

District-specific data from Figure 1a highlights Hailakandi in Assam with a staggering 64.7% and Bishnupur in Manipur at 59.9% as the areas experiencing the highest domestic violence rates between 2019 and 2021. Seven districts, including Bishnupur, Chandel, and Imphal West in Manipur and Kokhrajhar, Karimganj, Cachar, and Hailakandi in Assam, reported rates exceeding 50%, while six districts fell within the 40–50% range. Regionally, physical violence stands at 28%, with Manipur topping at 38.5%. Notably, Bishnupur in Manipur and Karimganj in Assam reported the highest rates, with four districts exceeding 50% and five falling within the 40–50% range, as depicted in Figure 2a. Emotional violence averages 11.7%, with Sikkim recording the highest at 14.6%. Assam's Hailakandi, Cachar, and Karimganj faced notably high rates, with four districts surpassing the 25% mark, as depicted in Figure 3a. Regarding sexual violence at 6.4%, Assam leads at 7%, with Karimganj in Assam and East Kameng in Arunachal Pradesh reporting the highest rates. Two districts reported rates above 15%, while fifteen districts fell within the 10% to 15% range, as illustrated in Figure 4a.

In Northeast India, Figure 1b shows the LISA cluster prevalence map for DV, presenting high-high clusters (10 districts), low-low clusters (14 districts), and low-high spatial outliers (2 districts) in 2019–2021. Moran's *I* value is 0.514, indicating positive spatial autocorrelation. Hotspot clusters were observed in Assam (Karimganj, Hailakandi, Cachar) and Manipur (Bishnupur, Thoubal, Imphal East, Imphal West, Churachandpur, Chandel, Senapati). Mizoram and Nagaland mostly fell into low-low clusters, except for specific districts.

Figure 2b displays the LISA cluster prevalence map for physical violence, illustrating high-high clusters (11 districts), low-low clusters (16 districts), and low-high spatial outliers (1 district) in 2019–2021. Moran's *I* is 0.550, suggesting positive spatial autocorrelation. Hotspot clusters were observed in Assam (Karimganj, Hailakandi, Chirang, Cachar) and Manipur (Bishnupur, Thoubal, Imphal East, Imphal West, Churachandpur, Chandel, Senapati). Certain districts in Meghalaya, Mizoram, and Nagaland mainly formed low-low clusters.

In Figure 3b, the LISA cluster prevalence map for emotional violence reveals high-high clusters (4 districts), low-low clusters (5 districts), and low-high spatial outliers (3 districts) in 2019–2021. Moran's *I* stand at 0.428, indicating positive spatial autocorrelation. High hotspot clusters were observed in Assam (Karimganj, Hailakandi, Cachar) and Meghalaya (Southwest Khasi hills). Figure 4b presents the LISA cluster prevalence map for sexual violence, displaying high-high clusters (8 districts), low-low clusters (11 districts), and low-high spatial outliers (3 districts) in 2019–2021. Moran's *I* is 0.392, signifying positive spatial autocorrelation. Notably, Assam (Karimganj, Hailakandi, Cachar) and specific districts in Arunachal Pradesh (Tawang, Kurung Kumey, Kra Daadi, Lower Subansiri, Papum pare) observe hot spot clusters.

In Figure 1c and Figure 2c, G_i^* hot spot clustering analysis of domestic and physical violence in each state of northeast India showcases similar trends. Most hotspot clusters occur in the districts Bishnupur, Thoubal, Imphal East, Imphal West, and Chandel district in Manipur, alongside Assam's Karimganj, Hailakandi, and Cachar districts. Cold spots are primarily observed in districts like Kiphere, Tuensang, Mokokchung, Zunheboto, Longleng, Kohima, and Dimapur in Nagaland. There are no identified hotspot clusters in Mizoram, Tripura, Nagaland, and Meghalaya, with the high-intensity hotspots mainly concentrated in the valley districts of Manipur.

Figure 3c displays the district-wise hotspot analysis of emotional violence in each state of Northeast India, emphasizing the major hotspot clusters found in Assam's Karimganj, Hailakandi, and Cachar districts. Similarly, Figure 4c reveals the district-wise hotspot analysis of sexual violence in each state of Northeast India, indicating the highest concentration of hotspot clusters in Assam's Karimganj, Hailakandi, and Cachar districts.

1. Domestic Violence

Figure 1a. District-wise prevalence in Northeast India.

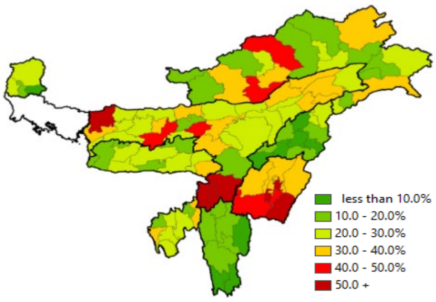


Figure 1b. District-wise LISA map in Northeast India.

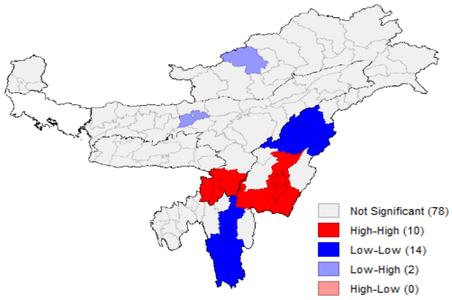
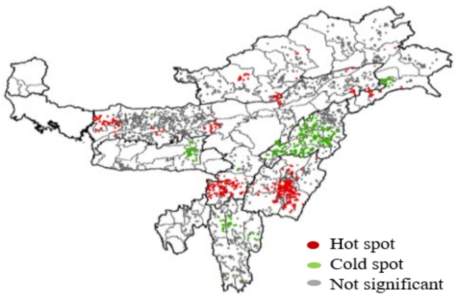


Figure 1c. G_i^* Hot spot clustering in Northeast India.



2. Physical Violence

Figure 2a. District-wise prevalence in Northeast India.

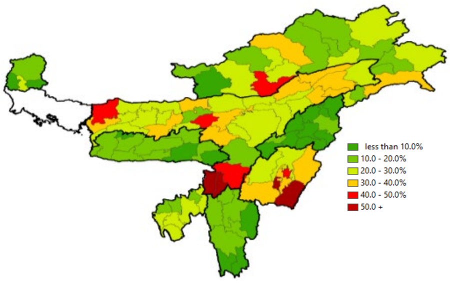


Figure 2b. District-wise LISA map in Northeast India.

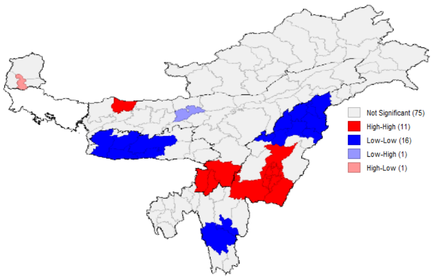
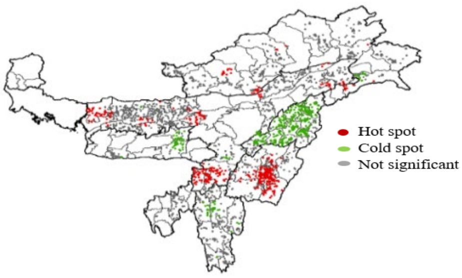


Figure 2c. G_i^* Hot spot clustering in Northeast India.



3. Emotional Violence

Figure 3a. District-wise prevalence in Northeast India.

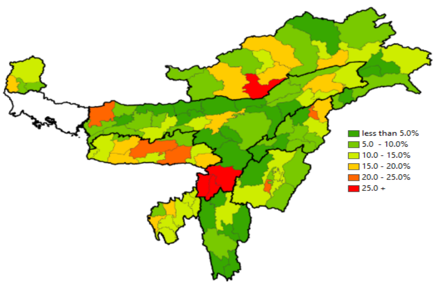


Figure 3b. District-wise LISA map in Northeast India.

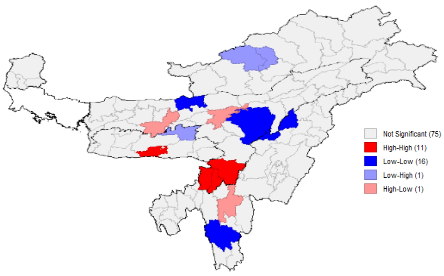
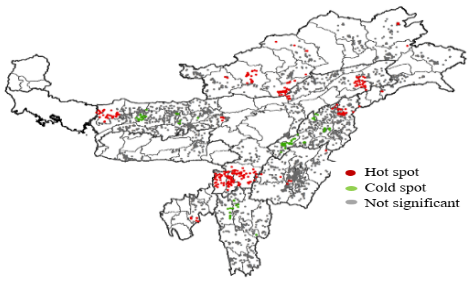


Figure 3c. G_i^* Hot spot clustering in Northeast India.



4. Sexual Violence

Figure 4a. District-wise prevalence in Northeast India.

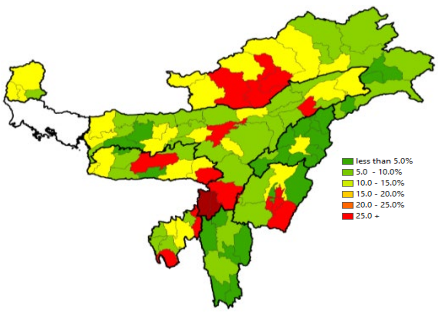


Figure 4b. District-wise LISA map in Northeast India.

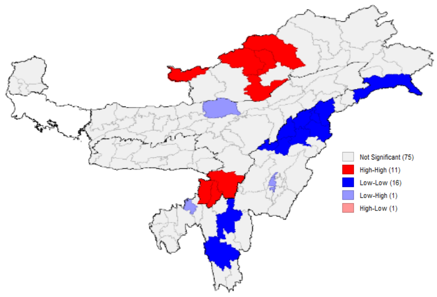
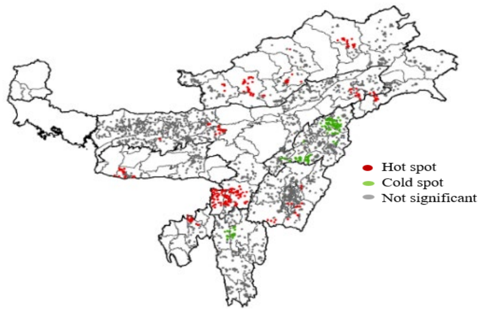


Figure 4c. G_i^* Hot spot clustering in Northeast India.



Note: Author's calculation using NFHS-5 data from the MoHFW, Government of India.

Comparison between OLS and GWR models

Unlike global models (OLS), geographically weighted regression (GWR) captures the dynamic spatial structure of the dataset by considering the non-stationary nature and underlying relationships of employed variables at the neighborhood level. The global model is initially essential for understanding the constant non-spatial effects of the covariates on dependent variables, which helps generate local models, such as GWR. This research applied GWR to the same independent variables to examine the non-stationary spatial effect on domestic violence prevalence at the local scale (district level) (Roy et al., 2024b). Table 2 outlines the differences in model specifications between OLS and GWR, highlighting the improvements seen when shifting from the global OLS to the local GWR. An effective way to gauge this transition's benefits is by comparing the AICc values between GWR and OLS.

Table 2: Comparison Between OLS and GWR Models

Value	OLS				GWR			
	Domestic Violence	Physical Violence	Emotional Violence	Sexual Violence	Domestic Violence	Physical Violence	Emotional Violence	Sexual Violence
AICc*	842.89	842.89	745.55	606.70	812.55	808.15	711.15	557.42
R^2 adjusted	0.21	0.22	0.10	0.09	0.52	0.59	0.39	0.58

Note: OLS = Ordinary Least Square, GWR = Geographically Weighted Regression. The author calculated this using NFHS-5 data from the MoHFW, Government of India.

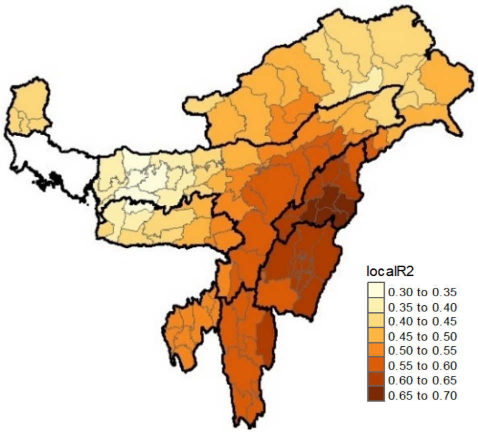
In the case of DV, the R^2 stands at 21% in OLS, whereas it substantially increases to 52% in GWR. Regarding physical violence, the adjusted R^2 value is 22% in OLS, notably rising to 59% in GWR. Similarly, for emotional and sexual violence, R^2 values are 10% and 9%, respectively, in OLS but demonstrate a significant increase to 39% and 58% in GWR. Moreover, the AICc value is considerably smaller in GWR than in OLS, indicating a more favorable performance. A smaller AICc value signifies a better fit of the model.

Local R^2 of geographically weighted regression

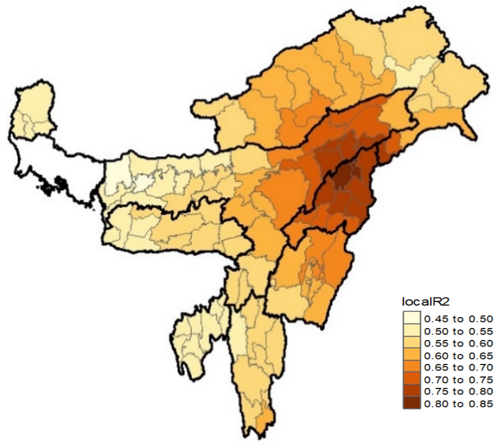
Geographically weighted regression (GWR) results have revealed spatial variations in the relationships between DV and its explanatory variables, showcasing differing significance and directions of these relationships across local areas. Figure 5 displays the local R^2 values concerning domestic violence and its various forms: physical, emotional, and sexual violence. The local R^2 values for DV were notably higher in the southern and eastern regions of the Northeast States. Specifically, in the southeastern districts of Nagaland, these values ranged between 0.65 and 0.70. The highest local R^2 values were observed for physical violence in Nagaland and its bordering districts in Assam and Arunachal Pradesh. Notably, Longleng and Mokokchung districts in Nagaland exhibited the highest R^2 values, ranging from 0.80 to 0.85. Concerning emotional violence, the local R^2 values were most prominent in the southern states of Northeast India. These values ranged from 0.45 to 0.50, encompassing various districts in Mizoram, Churachandpur, Tamenglong, Imphal East, Imphal West, Chandel, Bishnupur, and Thoubal in Manipur, as well as northern districts of Tripura, and Hailakandi, Karimganj, Cachar, and Dima Hasao districts in Assam. Lastly, in sexual violence, the highest local R^2 values were observed in the southern states of northeast India, particularly in Mizoram and Manipur. The R^2 values ranged from 0.65 to 0.70, covering the entire districts of Mizoram, Churachandpur, Tamenglong, Imphal East, Imphal West, Chandel, Bishnupur, and Thoubal in Manipur, along with the South Tripura district in Tripura.

Figure 5: Spatial Distribution of Local R^2 Values for GWR Analysis in the Different Forms of Violence in Northeast India.

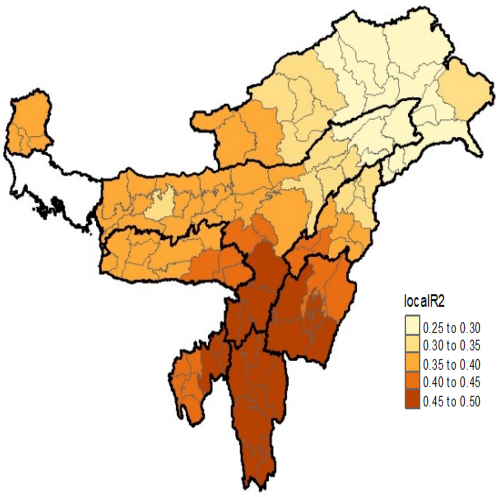
a) Domestic violence Local R^2 map



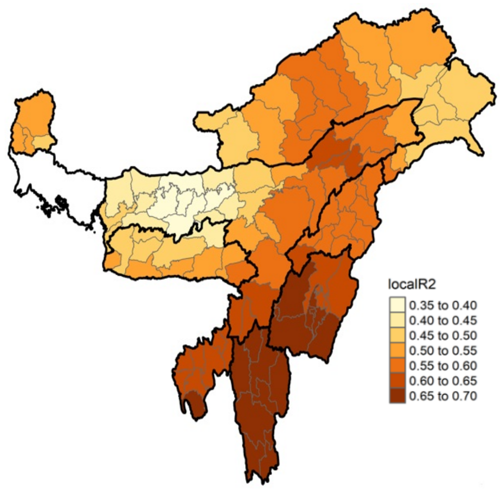
b) Physical violence Local R^2 map



c) Emotional violence Local R^2 map



d) Sexual violence Local R^2 map



Discussion

The mapped coefficients for each district of the Northeast indicated where the explanatory variables were effective predictors of domestic violence (DV) and where they were not. Figures 6.1, 6.2, 6.3, and 6.4 in the Appendix show that states or districts shaded in the darkest red represent the highest coefficient values. The larger the coefficient, the stronger the relationship. The proportion of women aged 35–49 had a positive relationship to domestic violence incidence; as the proportion coefficient of age increased, DV incidence also increased. All four dependent variables, such as DV, physical, emotional, and sexual violence, have the same proportion of coefficient value in women aged 35–49 in Northeast India. Likewise, the other independent variables such as the proportion of women who follow the Christian religion, working women, separated/ divorced women, the proportion of women whose husbands drink alcohol, poorest wealth quintile, and illiterate women have also equally likely the same proportion of coefficient values in DV, physical, emotional and sexual violence.

The proportion of women aged 35–49 strongly predicts demographics in several districts across Assam, Mizoram, Arunachal Pradesh, Manipur, and Nagaland, with coefficient values ranging from 54 to 70. However, Sikkim, Meghalaya, and Tripura have weaker relationships, with coefficients below 50. In most districts of Nagaland, Mizoram, specific areas in Manipur (Churachandpur, Tamenglong, and Ukhrul), and certain regions in Meghalaya (South Garo Hills, South West Khasi Hills, and West Khasi Hills), the Christian population is notably higher, with coefficient values between 95 and 100. Conversely, Arunachal Pradesh, Sikkim, Assam, and Tripura display remarkably low coefficients regarding the Christian religion.

In several districts of Meghalaya, Nagaland, Manipur, and certain parts of Arunachal Pradesh, the coefficient for working women is high, ranging between 44 and 66. These areas include specific districts like Anjaw, Dibang Valley, Lower Dibang Valley, Upper Siang, Upper Subansiri, and Kamle in Arunachal Pradesh; Longleng, Wokha, Kohima, ZunHebeto, Phek in Nagaland; Ribhoi, West Jaintia Hills, East Jaintia Hills, West Khasi Hills, South West Khasi Hills in Meghalaya; and Bishnupur, Thoubal, Imphal West in Manipur. Conversely, Tripura, Mizoram, Assam, and Sikkim show lower coefficient values for the proportion of working women. In most districts of Mizoram, Meghalaya, specific parts of Manipur (Churachandpur, Ukhrul), Dimapur in Nagaland, South and West Sikkim, Cachar and Sivasagar in Assam, and Kurung Kumey and Papum Pare in Arunachal Pradesh, divorced/separated couples they had higher proportions, ranging from 4 to 15 in coefficient values. Tripura had lower rates of separated or divorced individuals, with coefficient values less than 3.90 (Figure 6.1).

The easternmost parts of the Northeast, excluding Nagaland, display stronger coefficients for husbands' alcohol use. Districts like West Sikkim, Khowai in Tripura, West Jaintia Hills, and West Garo Hills in Meghalaya, and various districts in Assam, Arunachal Pradesh, and Manipur show high coefficients (50 to 70) for husbands’ alcohol consumption. In contrast, Mizoram and Nagaland have notably lower coefficients for husbands’ alcohol use in Northeast India.

The central regions of the Northeast exhibit stronger coefficients for the poorest wealth quintile. High coefficient values, ranging from 45 to 63, are observed in most districts of Assam, Tripura, Meghalaya, certain northeastern districts in Nagaland, and Ukhrul district in Manipur. Conversely, Mizoram, Arunachal Pradesh, and Sikkim show a low coefficient value. Additionally, northeast India's northern and central parts demonstrate a higher proportion of illiterate women. Manipur, Sikkim, Mizoram, Nagaland (excluding Dhalai district), and Tripura exhibit low coefficients for illiterate women.

Our findings are consistent with other studies conducted in India and abroad; spatial clustering of DV was observed, even within districts of a state (Roy & Chowdhury, 2023a; Seid et al., 2021). Researchers have also reported that the risk factors for spousal violence include younger age married women, drinking husbands, and low-income and non-working women (Babu & Kar, 2009; Haobijam & Singh, 2022; Koenig et al., 2006). In another study conducted in Gujarat, protective factors for DV were explored, and women with higher socioeconomic status and economic independence were associated with a lower risk of experiencing DV (Visaria, 2000). However, our data was based on 2019–2021, and studies have reported a significant increase in the country’s DV rate during the COVID-19 pandemic. The factors for such a high prevalence rate include travel restrictions, home containment, increased alcohol use by partners, and a higher rate of unemployment, which aggravated DV at an alarming rate during the COVID-19 pandemic (Krishnakumar & Verma, 2021; Maji et al., 2022). These factors could have impacted our findings as well; however, being a secondary data-based analysis, such issues are beyond the scope of our study.

In our study, GWR results highlight localized variations in the relationship between DV and its explanatory variables, with significance and direction varying across different spatial regions. Except for women's age group and husbands' use of alcohol, other factors show inconsistent predictive strengths across districts of Northeast India. This may be because Northeast India is diverse and has many socio-culturally unique groups. This underscores the need to conduct more extensive studies in this region. Further, efforts are needed to develop appropriate district-based intervention programs rather than adopting a generic and general policy.

Limitation and strengths

The data used in this analysis is from nationally representative NFHS-5 surveys, which are exhaustive and comprehensive. However, the limitations of the original data apply to the present study as well. The sample consists of a diverse population from all corners of the country, including women from all backgrounds, religions, regions, castes, cultures, creeds, and socioeconomic statuses. The study does not include contributing factors that might have led to the high prevalence of DV, such as the unemployment rate (Seid et al., 2021), psycho-social factors, and environmental factors.

The GWR offers the advantage of generating local parameter coefficients for each data point and enabling model diagnostics. This capability allows visualization and interpretation of spatial variations within the dataset. However, handling the substantial amount of spatial data produced by GWR presents challenges, particularly in simultaneously presenting parameter estimates and their associated significance, such as t-values, for accurate interpretation. In our mapping process, we illustrated spatial distributions of local R^2 and GWR parameter coefficients while applying a 95% significance threshold to mask points where the relationship between the dependent variable and predictor was insignificant.

Conclusion

The comprehensive analysis of domestic violence (DV) across Northeast India between 2019 and 2021 reveals regional disparities and varying intensities across different forms of violence. Manipur recorded the highest overall prevalence at 41.5%, followed by Assam, Arunachal Pradesh, Tripura, Meghalaya, Sikkim, Mizoram, and Nagaland. Districts like Hailakandi in Assam and Bishnupur in Manipur stood out with staggering rates exceeding 50%. Regional disparities were evident in the prevalence of physical, emotional, and sexual violence, with Manipur, Assam, and specific districts in Arunachal Pradesh and Nagaland frequently highlighted as hotspots. Additionally, areas in Mizoram, Nagaland, Meghalaya, and Tripura formed low-intensity clusters for violence incidents, contrasting with the high-intensity hotspots concentrated in Manipur valley districts.

The transition from OLS to GWR revealed significant enhancements in the R^2 values for all forms of violence, demonstrating an increase from 21% to 52% for domestic violence, 22% to 59% for physical violence, 10% to 39% for emotional violence, and 9% to 58% for sexual violence. Moreover, the smaller AICc values in GWR indicated a superior fit of the model. The local R^2 values derived from GWR unveiled spatial variations across the Northeast, emphasizing higher predictive strengths in specific districts, notably in southern and eastern regions for DV and distinct clusters in Nagaland and bordering districts for physical violence. Additionally, the relationship between explanatory variables and violence incidents varied across local areas, with certain factors like the proportion of women aged 35–49 showing consistent predictive strengths across all forms of violence, while others exhibited disparities across different regions. These findings emphasize the need for targeted interventions and region-specific approaches to address and prevent DV effectively across Northeast India.

The prevention and control of various forms of DV is a shared responsibility; thereby, all parties, including governmental organizations, non-governmental organizations, the scientific community, community leaders, and every individual, should be involved. The government should prioritize addressing key factors contributing to IPV in hotspot areas. The factors include high levels of alcohol consumption, as well as the vulnerabilities faced by working women and illiterate women. By focusing on these issues, intervention activities can be more effective in reducing IPV and creating safer environments for women. Additionally, the scientific community needs to uncover the hidden realities of DV by conducting thorough research (Seid et al., 2021).

While addressing the issue of DV, it is essential to focus on the core underlying factors such as alcohol use by husbands, women's literacy and working status, and specific age groups of women. Mass awareness campaigns, psycho-education, and community outreach programs through women's organizations and local clubs can be helpful, especially in regions with hotspot clusters. Religious and community leaders, health care workers, and officials of law-enforcing agencies should be sensitized to the issue as they can play a pivotal role in reducing the prevalence of DV.

Acknowledgments

The authors thank the Journal of Population and Social Studies (JPSS), the Institute for Population and Social Research (IPSR), and Mahidol University for their support and two anonymous reviewers for their helpful comments.

References

Abramsky, T., Watts, C. H., Garcia-Moreno, C., Devries, K., Kiss, L., Ellsberg, M., Jansen, H. A. F. M., & Heise, L. (2011). What factors are associated with recent intimate partner violence? Findings from the WHO multi-country study on women's health and domestic violence. *BMC Public Health, 11*, Article 109. https://doi.org/10.1186/1471-2458-11-109

Babu, B. V., & Kar, S. K. (2009). Domestic violence against women in eastern India: A population-based study on prevalence and related issues. *BMC Public Health, 9*, Article 129. https://doi.org/10.1186/1471-2458-9-129

Baruah, A., & Baruah, S. (2021). The 'Golden Triangle'–A hub of narcotics trade and its repercussions for North-East India vis-à-vis the Act East Policy–An analysis. *The Indian Police Journal, 68*(1), 160–173. https://bprd.nic.in/uploads/pdf/202106010319281112841IPJJanToMar2021.pdf

Basumatary, N., & Panda, B. (2020). A review of institutional and developmental issues in North-East India. *Indian Journal of Public Administration, 66*(2), 206–218. https://doi.org/10.1177/0019556120916884

Dahlberg, L. L., & Krug, E. G. (2006). Violence: A global public health problem. *Ciência & Saúde Coletiva, 11*(2), 277–292. https://doi.org/10.1590/S1413-81232006000200007

Devries, K., Watts, C., Yoshihama, M., Kiss, L., Schraiber, L. B., Deyessa, N., Heise, L., Durand, J., Mbwambo, J., Janssen, H., Berhane, Y., Ellsberg, M., & Garcia-Moreno, C. (2011). Violence against women is strongly associated with suicide attempts: Evidence from the WHO multi-country study on women's health and domestic violence against women. *Social Science & Medicine, 73*(1), 79–86. https://doi.org/10.1016/j.socscimed.2011.05.006

Gazette of India. (2005, September 13). *The Protection of Women from Domestic Violence Act, 2005 (Act No. 43 of 2005).* https://www.indiacode.nic.in/bitstream/123456789/15436/1/protection_of_women_from_domestic_violence_act%2C_2005.pdf

Haobijam, S., & Singh, K. A. (2022). Socioeconomic determinants of domestic violence in Northeast India: Evidence from the National Family Health Survey (NFHS-4). *Journal of Interpersonal Violence, 37*(15–16), NP13162–NP13181. https://doi.org/10.1177/08862605211005133

Haobijam, S., & Singh, K. A. (2023). Risk factors and perpetrators of domestic violence during any pregnancy experienced by women in North-East India. *Partner Abuse, 14*(3), 366–381. https://doi.org/10.1891/PA-2022-0056

International Institute for Population Sciences (IIPS) & ICF. (2017, December). *India National Family Health Survey (NFHS-4), 2015-16.* http://dhsprogram.com/pubs/pdf/FR339/FR339.pdf

International Institute for Population Sciences (IIPS) & ICF. (2022, March). *National Family Health Survey (NFHS-5), 2019-21: India: Volume II.* https://dhsprogram.com/pubs/pdf/FR375/FR375_II.pdf

Jamshidimanesh, M., Soleymani, M., Ebrahimi, E., & Hosseini, F. (2013). Domestic violence against pregnant women in Iran. *Journal of Family and Reproductive Health, 7*(1), 7–10. https://jfrh.tums.ac.ir/index.php/jfrh/article/view/306

Kalokhe, A., del Rio, C., Dunkle, K., Stephenson, R., Metheny, N., Paranjape, A., & Sahay, S. (2017). Domestic violence against women in India: A systematic review of a decade of quantitative studies. *Global Public Health, 12*(4), 498–513. https://doi.org/10.1080/17441692.2015.1119293

Koenig, M. A., Stephenson, R., Ahmed, S., Jejeebhoy, S. J., & Campbell, J. (2006). Individual and contextual determinants of domestic violence in North India. *American Journal of Public Health, 96*(1), 132–138. https://doi.org/10.2105/AJPH.2004.050872

Krishnakumar, A., & Verma, S. (2021). Understanding domestic violence in India during COVID-19: A routine activity approach. *Asian Journal of Criminology, 16*(1), 19–35. https://doi.org/10.1007/s11417-020-09340-1

Lahiri, P. (2017, April 21). Lack of development in the northeastern part of India: A historical legacy. *Concept Research Foundation.* https://conceptresearchfoundation.com/2017/04/21/lack-of-development-in-the-north-eastern-part-of-india-%E2%80%95-a-historical-legacy/

Maji, S., Bansod, S., & Singh, T. (2022). Domestic violence during COVID-19 pandemic: The case for Indian women. *Journal of Community and Applied Social Psychology*, *32*(3), 374–381. https://doi.org/10.1002/casp.2501

Pallitto, C. C., García-Moreno, C., Jansen, H. A. F. M., Heise, L., Ellsberg, M., & Watts, C. (2013). Intimate partner violence, abortion, and unintended pregnancy: Results from the WHO multi-country study on women's health and domestic violence. *International Journal of Gynecology & Obstetrics, 120*(1), 3–9. https://doi.org/10.1016/j.ijgo.2012.07.003

Potter, L. C., Morris, R., Hegarty, K., García-Moreno, C., & Feder, G. (2021). Categories and health impacts of intimate partner violence in the World Health Organization multi-country study on women's health and domestic violence. *International Journal of Epidemiology, 50*(2), 652–662. https://doi.org/10.1093/ije/dyaa220

Roy, S., & Chowdhury, I. R. (2023a). Geography of crime against women in West Bengal, India: Identifying spatio-temporal dynamics and hotspots. *GeoJournal, 88*(6), 5863–5895. https://doi.org/10.1007/s10708-023-10946-y

Roy, S., & Chowdhury, I. R. (2023b). Three decades of GIS application in spatial crime analysis: Present global status and emerging trends. *Professional Geographer, 75*(6), 882–904. https://doi.org/10.1080/00330124.2023.2223250

Roy, S., Majumder, S., Bose, A., & Chowdhury, I. R. (2023). Does geographical heterogeneity influence urban quality of life? A case of a densely populated Indian city. *Papers in Applied Geography, 9*(4), 395–424. https://doi.org/10.1080/23754931.2023.2225541

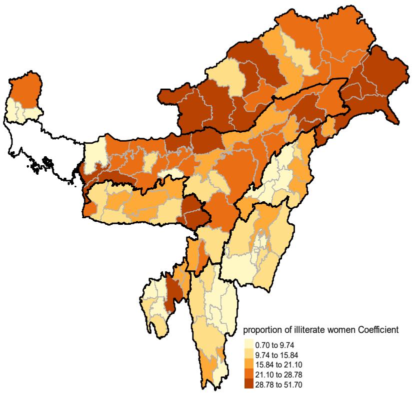
Roy, S., Majumder, S., Bose, A., & Chowdhury, I. R. (2024a). Spatial heterogeneity in urban household living conditions: A GIS-based spatial analysis. *Annals of GIS, 30*(1), 81–104.

https://doi.org/10.1080/19475683.2024.2304194
Roy, S., Majumder, S., Bose, A., & Chowdhury, I. R. (2024b). Understanding geographical variability of women's alcohol prevalence in India using spatial regression models: Evidence from a national representative survey. *Regional Science Policy and Practice, 16*(6), Article 100040. https://doi.org/10.1016/j.rspp.2024.100040
Seid, E., Melese, T., & Alemu, K. (2021). Spatial distribution and predictors of domestic violence against women: Evidence from analysis of the Ethiopian Demographic Health Survey 2016. *BMC Women's Health, 21*(1), Article 334. https://doi.org/10.1186/s12905-021-01465-4
Shidhaye, P. R., & Giri, P. A. (2014). Domestic violence during pregnancy: A neglected and unseen public health problem. *International Journal of Community Medicine and Public Health, 1*(1), 1–3. https://doi.org/10.5455/2394-6040.ijcmph20141101
Singh, A. K., & Suman, S. (2016). Immigrants issue in governance of North-East India: A fait accompli and humanitarian compulsions. *Indian Journal of Public Administration, 62*(3), 572–580. https://doi.org/10.1177/0019556120160317
Singh, K. J., Haobijam, N., Nair, S., Devi, A. S., Singh, S. R., Hijam, M., Alee, N. T., Sharma, S., Deepani, V., Singh, L., & Rao, M. V. V. (2021). Smokeless tobacco use among women in northeastern states, India: A study of spatial clustering and its determinants using National Family Health Survey-4 data. *Clinical Epidemiology and Global Health, 12*, Article 100840. https://doi.org/10.1016/j.cegh.2021.100840
Visaria, L. (2000). Violence against women. *Economic and Political Weekly, 94*(7), 157–158. https://doi.org/10.48029/nji.2003.lxxxxiv703
World Health Organization (WHO). (2005, February 9). *WHO multi-country study on women's health and domestic violence against women: Summary report. Initial results on prevalence, health outcomes and women's responses.* World Health Organization. https://www.who.int/publications/i/item/9241593512
World Health Organization (WHO). (2021, March 9). *Violence against women prevalence estimates, 2018: Global, regional, and national prevalence estimates for intimate partner violence against women and global and regional prevalence estimates for non-partner sexual violence against women.* https://www.who.int/publications/i/item/9789240022256

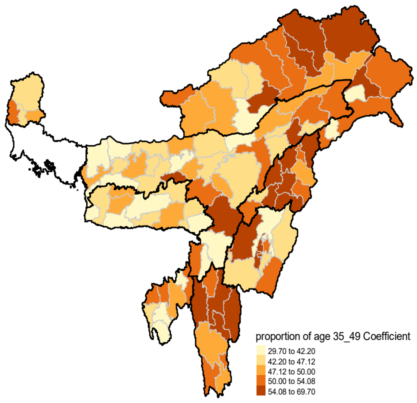
Appendix

Figure 6.1: Spatial Distribution of Local Regression Coefficients of Domestic Violence With Independent Variables

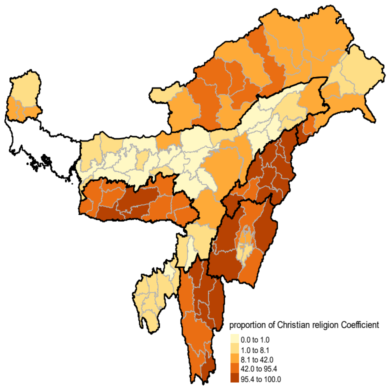
a. Illiterate women



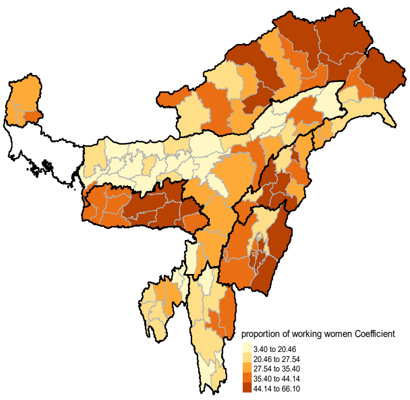
b. Age 35–49



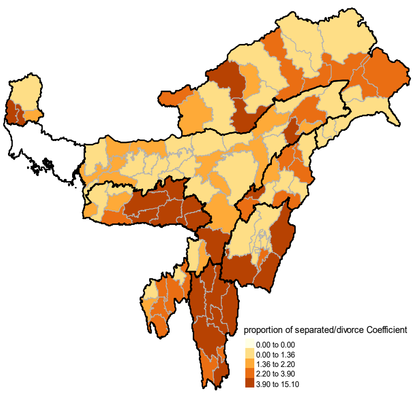
c. Christian religion



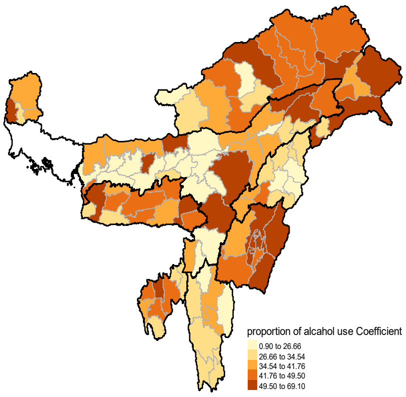
d. Working women



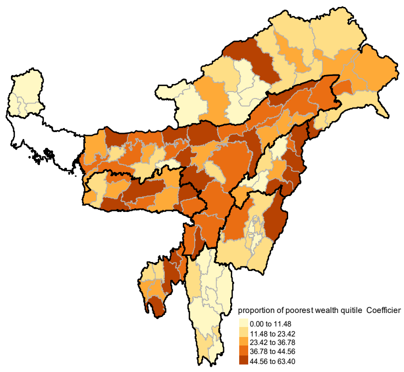
e. Separated/Divorce



f. Alcohol use

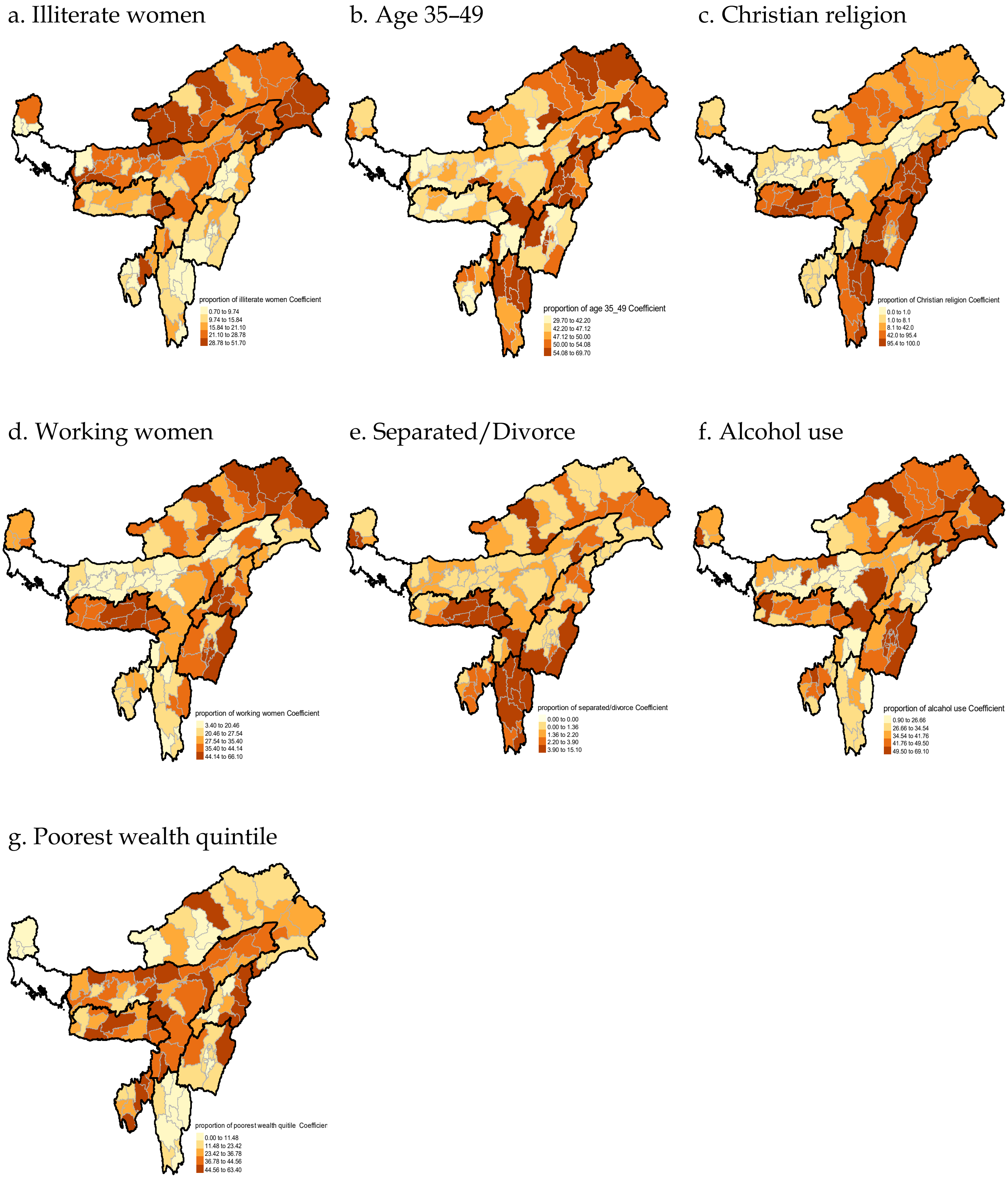


g. Poorest wealth quintile



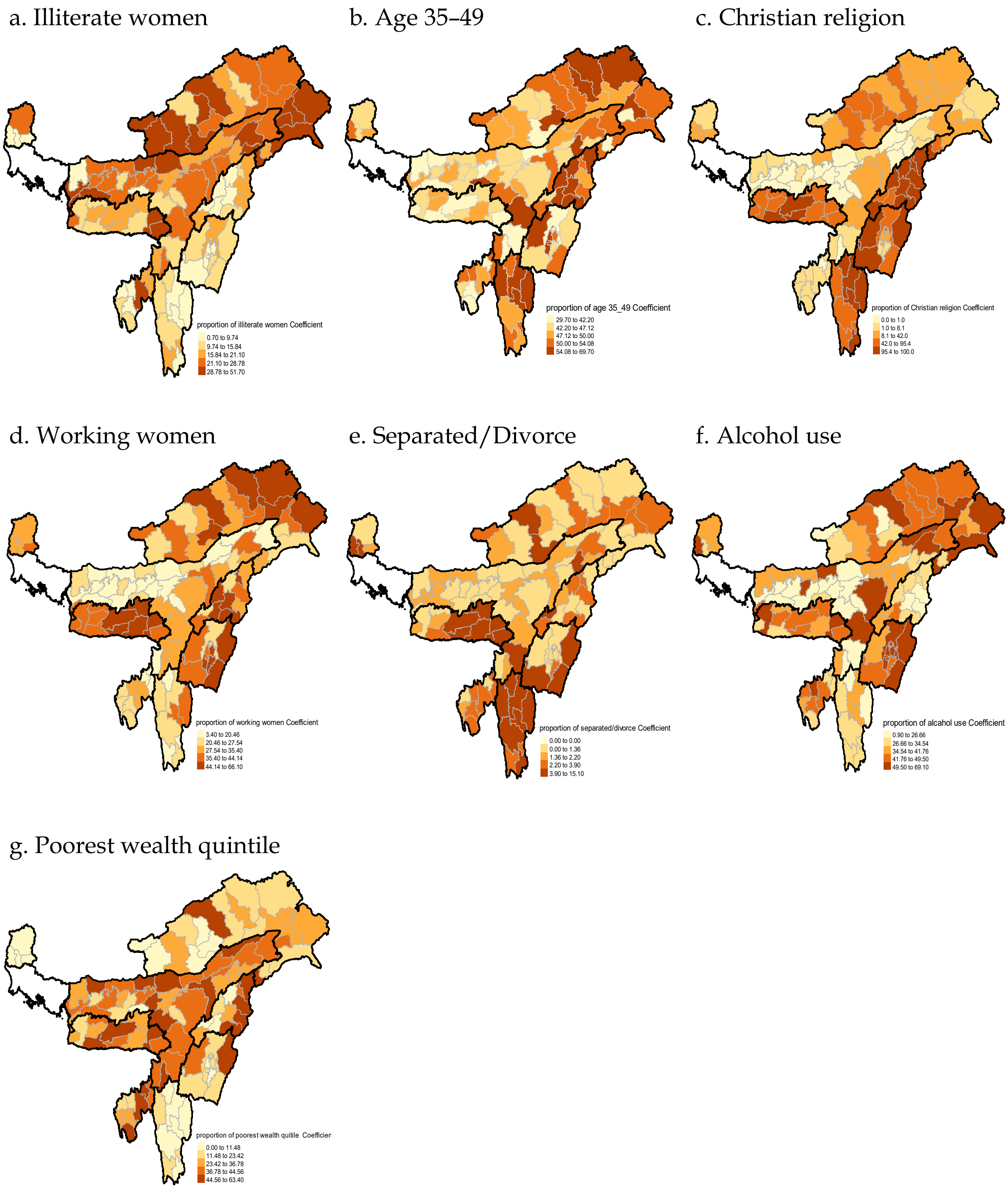
Note: Author's calculation using NFHS-5 data from the MoHFW, Government of India

Figure 6.2: Spatial Distribution of Local Regression Coefficients of Physical Violence With Independent Variables



Note: Author's calculation using NFHS-5 data from the MoHFW, Government of India

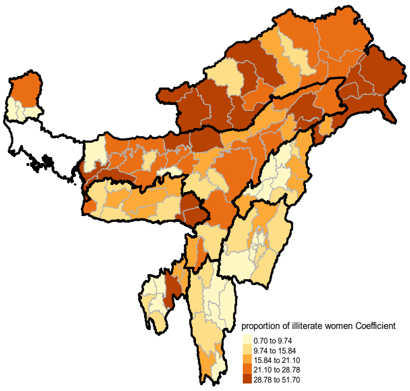
Figure 6.3: Spatial Distribution of Local Regression Coefficients of Emotional Violence With Independent Variables



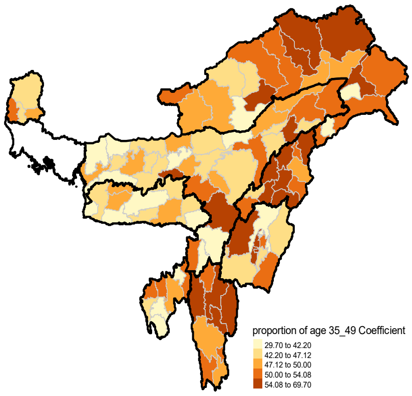
Note: Author's calculation using NFHS-5 data from the MoHFW, Government of India

Figure 6.4: Spatial Distribution of Local Regression Coefficients of Sexual Violence With Independent Variables

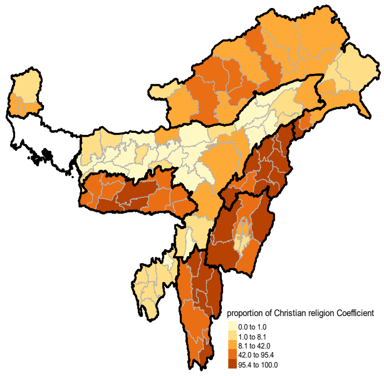
a. Illiterate women



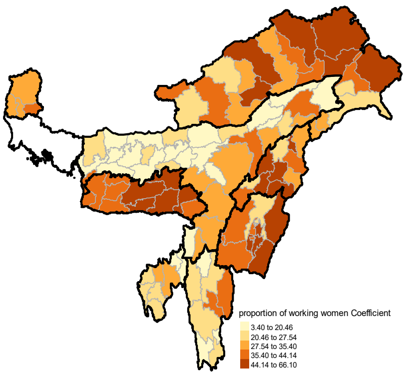
b. Age 35–49



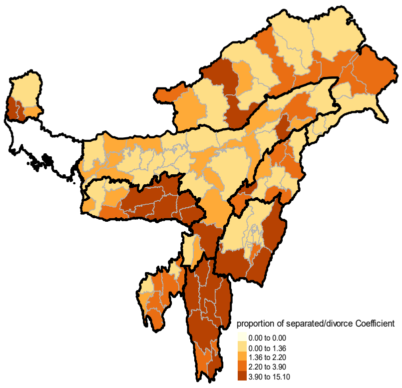
c. Christian religion



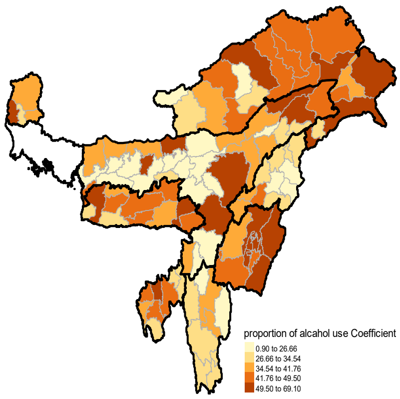
d. Working women



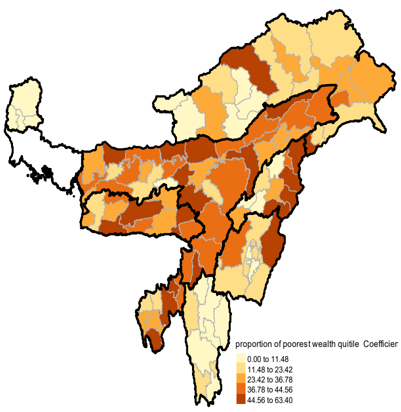
e. Separated/Divorce



f. Alcohol use



g. Poorest wealth quintile



Note: Author's calculation using NFHS-5 data from the MoHFW, Government of India