Leveraging Spatial Techniques in Observational Public Health Research

Explore how cutting-edge spatial techniques are revolutionising public health research, transforming data into actionable insights for effective policy-making.

Dr. Khangembam Jitenkumar Singh

Scientist E
Indian Council of Medical Research
New Delhi

02

Why Space/Location Matters in Public Health

- Location plays a pivotal role in understanding disease distribution, health disparities, and environmental influences.
- By mapping health outcomes using spatial techniques, we can identify areas of concern and target interventions more effectively.

Understanding **where** health outcomes occur,
helps us understand, **why** they occur.

03

Spatial Techniques in Observational Public Health Research

Spatial techniques are crucial for effective public health interventions:

- Analyze geographic patterns and clusters of disease incidence and screening uptake.
- Identify high-risk areas and vulnerable populations for targeted action.
- Support efficient resource allocation for public health interventions.

04

Foundations of Spatial Analysis

Geocoding

The process of converting addresses or locations into geographic coordinates, enabling data to be plotted on a map.

- Cancer registry data (2010-2024) geocoded at district and village levels.

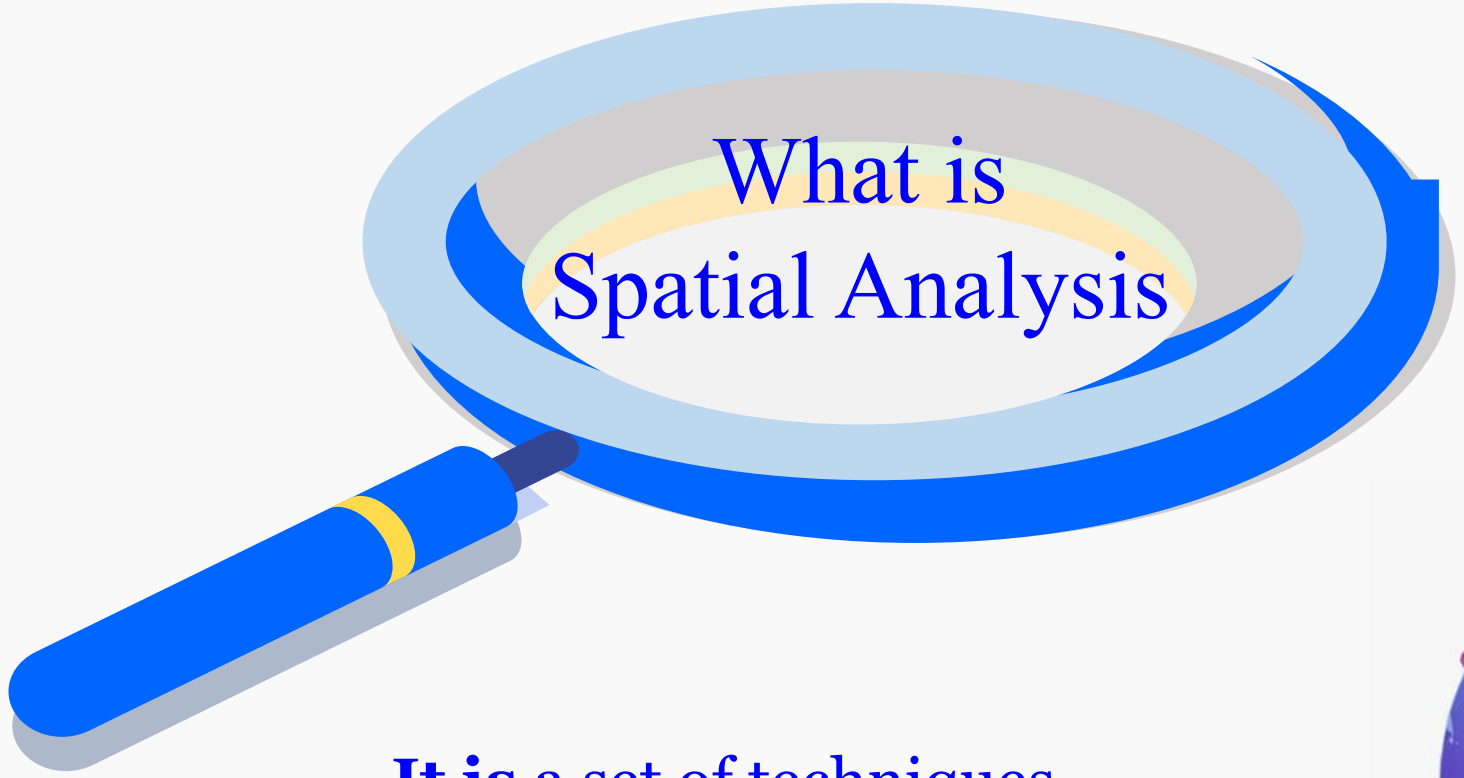
Geographic Information Systems (GIS)

Powerful software tools that capture, store, analyse, and present spatial or geographic data.

- Allows to visualise patterns, relationships, and trends.

05

What is Spatial Analysis



It is a set of techniques
to examine the *locations*,
patterns, *relationships*,
and *trends* of spatial data.

06

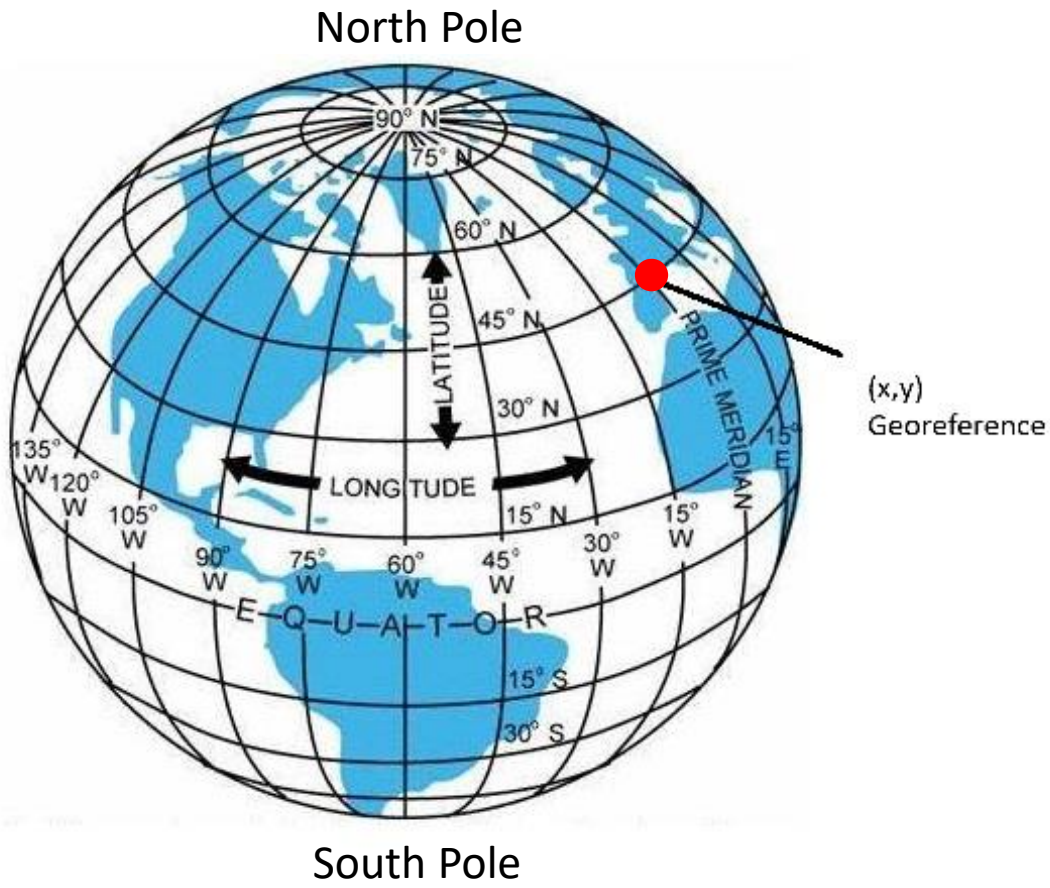
🔍 Key Concepts in Spatial Analysis

Concept	Meaning
Location	Where things happen (e.g., coordinates, villages, districts, states)
Distance	How far apart things are
Spatial patterns	Are features clustered, random, or evenly spread out?
Spatial relationships	Do nearby areas/locations influence each other?
Spatial dependence/ autocorrelation	Nearby areas/locations are more similar than distant ones (e.g., high c-section rates in adjacent districts)

What is a GPS?

07

- GPS - Global Positioning System
- Multiple satellites in space



X= easting = Longitude = horizontal line

Y= northing = Latitude = vertical line

08

Spatial Data Types

Types of spatial data:

Vector data

- ***Point*** : a single point location, such as a GPS reading or a geocoded address
- ***Line*** : a set of ordered points, connected by straight line segments
- ***Polygon***: an area, marked by one or more enclosing lines

Raster data

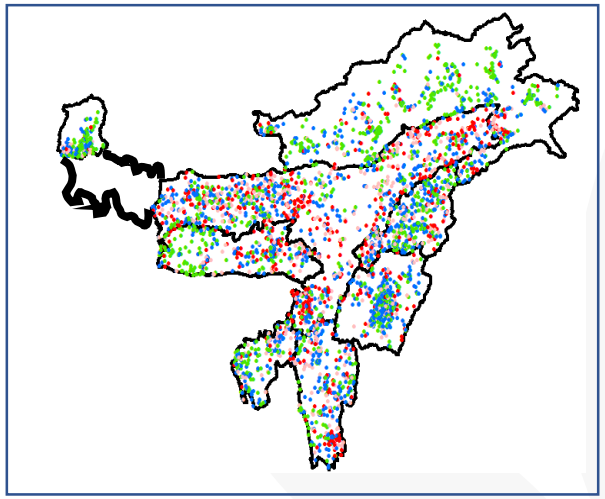
- ***Grid*** : a collection of points or rectangular cells, organized in a regular lattice

The first three are vector data models and fourth data model is a raster data model.

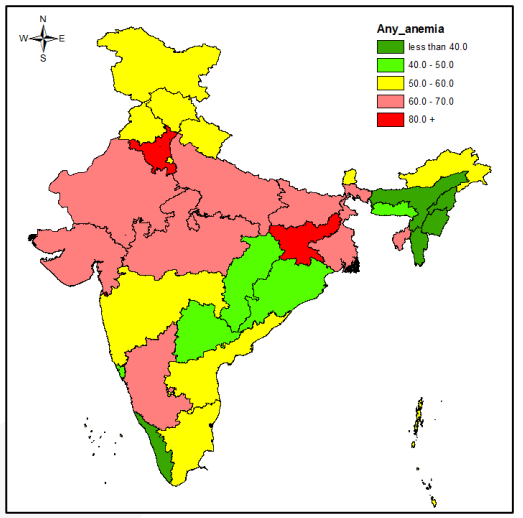
All spatial data consist of positional information, answering the question **“where is it?”.**

Example of geospatial / spatial data

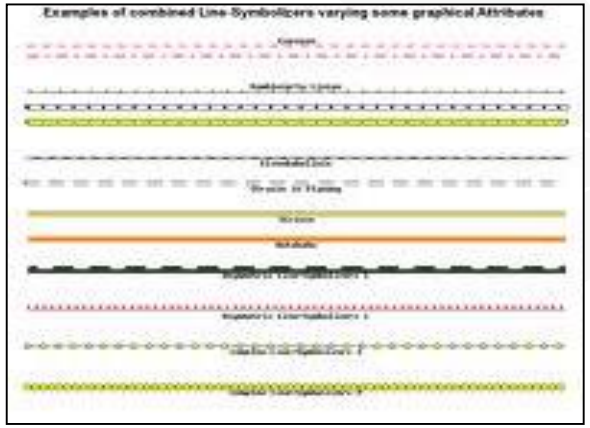
09



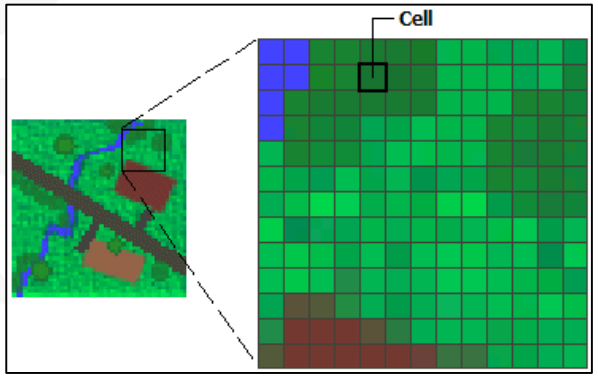
Point



Polygon



Line



Grid

Measuring Space

Not easy!

Everything is related to everything else,
but near things are more related than
distant things.

(*Tobler's First Law of Geography, 1970*)

→ Spatial dependence → spatial autocorrelation

Types of spatial analysis

1. **Descriptive Mapping**

   Just visualizing the data (e.g., disease prevalence on a map).

2. **Hotspot/Cluster Analysis**

   Identifies areas with unusually high or low values (e.g., hotspots of cervical cancer cases).

3. **Spatial Autocorrelation**

   Tests whether values are spatially clustered (e.g., Moran's I, Geary's C, LISA).

4. **Spatial Regression**

   Models outcomes while accounting for spatial dependence (e.g., spatial lag/error models).

5. **Geographically Weighted Regression (GWR)**

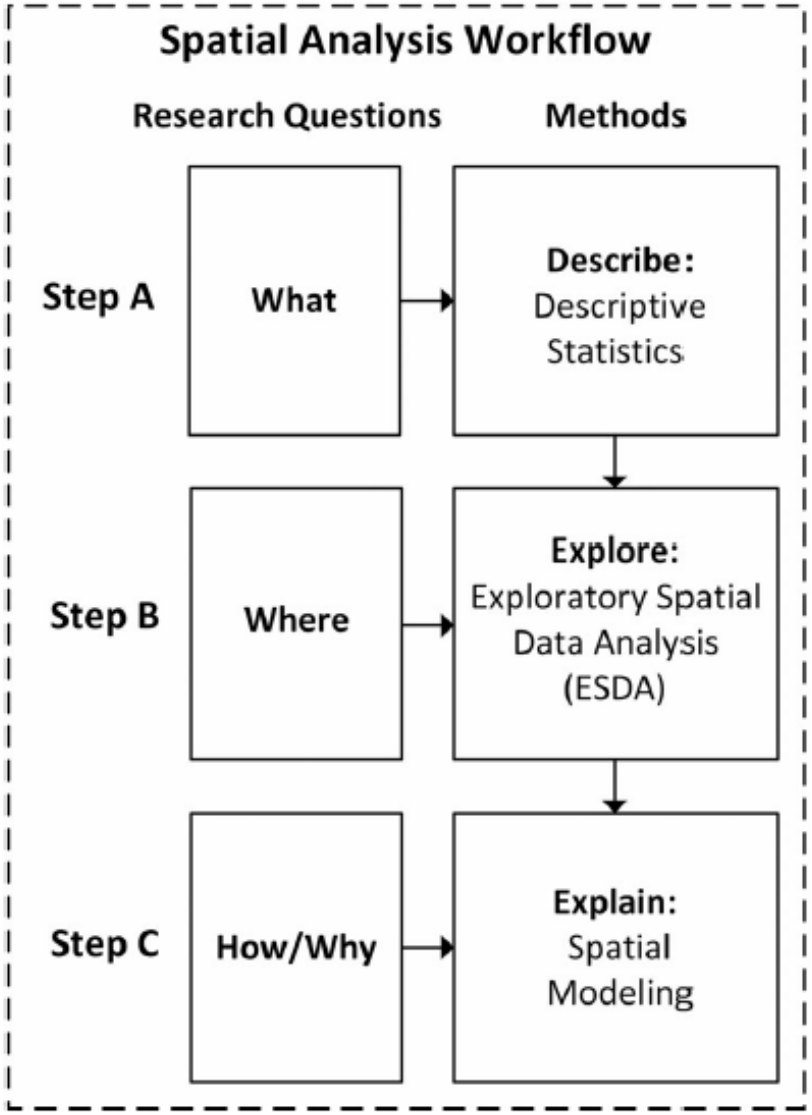
   Assesses how relationships vary across space.

6. **Interpolation**

   Estimates values at unsampled locations (e.g., kriging, Inverse Distance Weighting).

❖ Spatial analysis of point data, areal data, geostatistical data and Spatial modelling

Spatial Analysis Workflow



Descriptive statistics- summarize data characteristics and provide a useful understanding about the distribution of the values, their range and the presence of outliers.

ESDA is applied to explore and map data, locate outliers, test underlying assumptions or identify trends and associations among them (autocorrelation presence or spatial clustering).

Mostly answer "where?" questions, such as where are the areas with low/high C-section birth, is there any spatial clustering in the distribution, where is it located, and where are the hotspots?

Here attempt to answer "why?/how?" questions. These methods not only identify associations but also attempt to (a) unveil relations that explain why something happens and (b) trace the drivers behind a change.

Spatial Epidemiology

Mainly concerned between the place and health the population with the two fundamental questions:

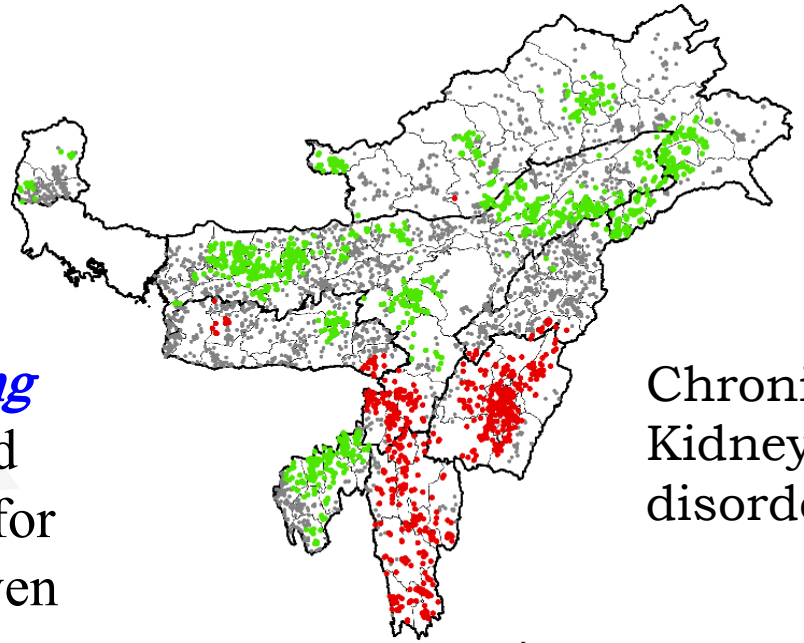
- where and when diseases tend to occur?
- where does such patterns exist and how/why?

Overview of Spatial Techniques

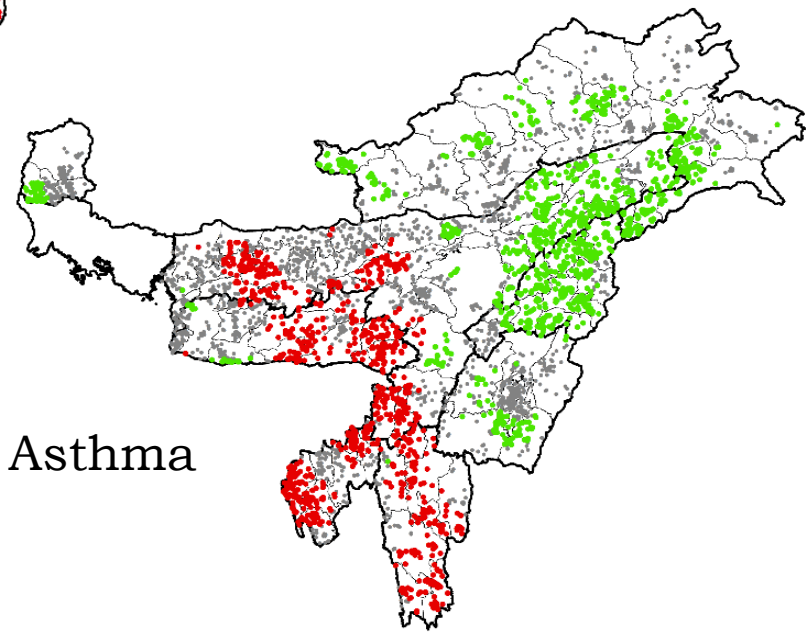
Method	Purpose	Tools
Spatial Autocorrelation (Moran's, Geary's C)	Detect clustering	R, STATA, GeoDa
Hotspot Analysis (LISA, Getis-Ord G_i^*)	Identify hot/cold spots	R, ArcGIS
Spatial Regression (SAR, SEM, SLM)	Adjust for spatial dependence	R, Stata, GeoDa
Geographically Weighted Regression (GWR, MGWR)	Local variability in associations	R, Stata, GWR4
Spatial Scan Statistics	Detect space-time clusters	SaTScan
Spatio-temporal	identify clusters, trends, and relationships across both space and time	R

The main advantage of spatial analysis is the ability to reveal patterns in data that had not previously been defined or even observed.

For example, using spatial analysis techniques, one might identify the ***clustering of a disease occurrence*** and then develop mechanisms for preventing expansion or even eliminating it.



Chronic Kidney disorder



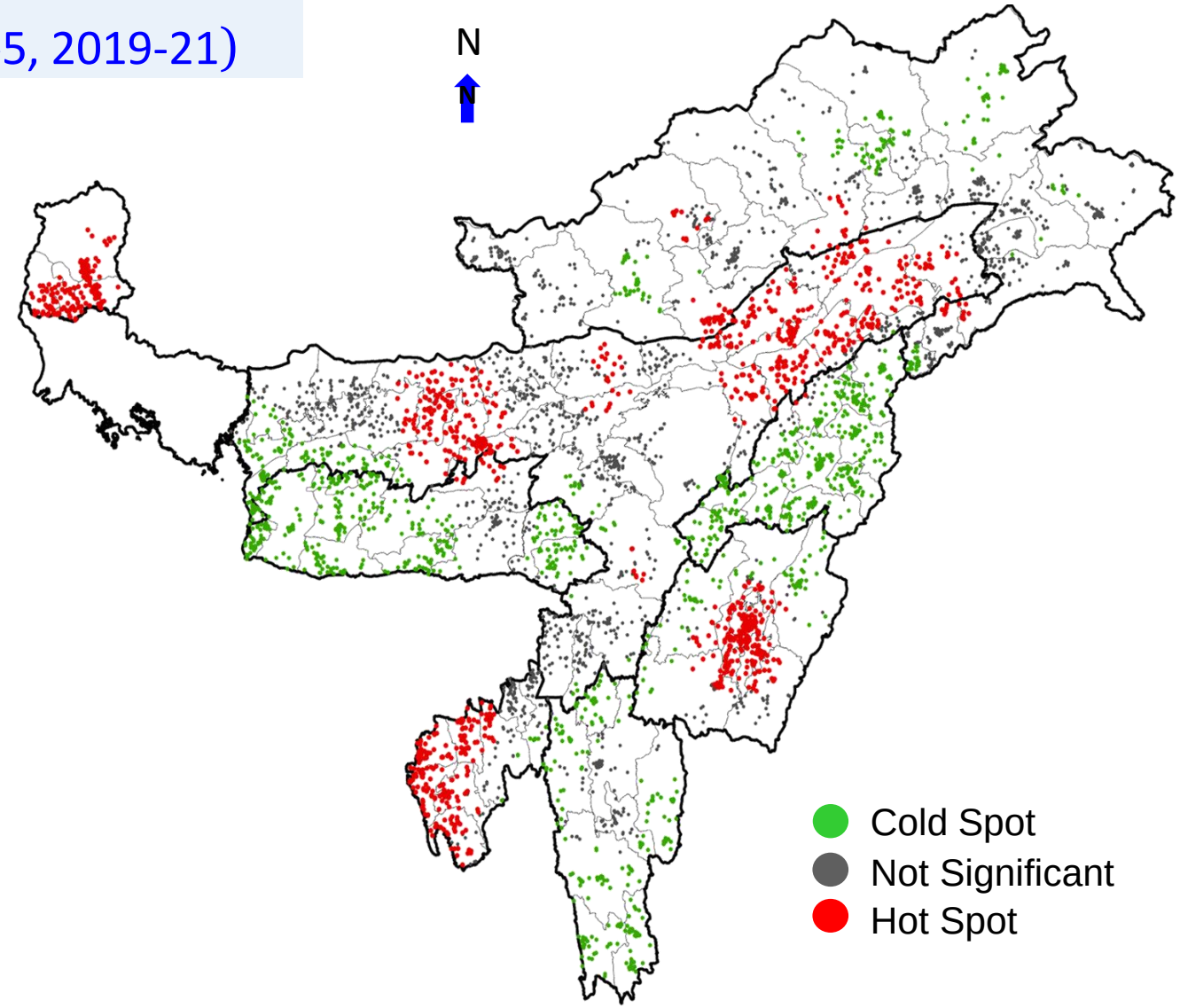
Asthma

In this respect, spatial analysis leads to better decision making and spatial planning.

Hot spot clustering of C-section delivery, Northeastern states (NFHS-5, 2019-21)

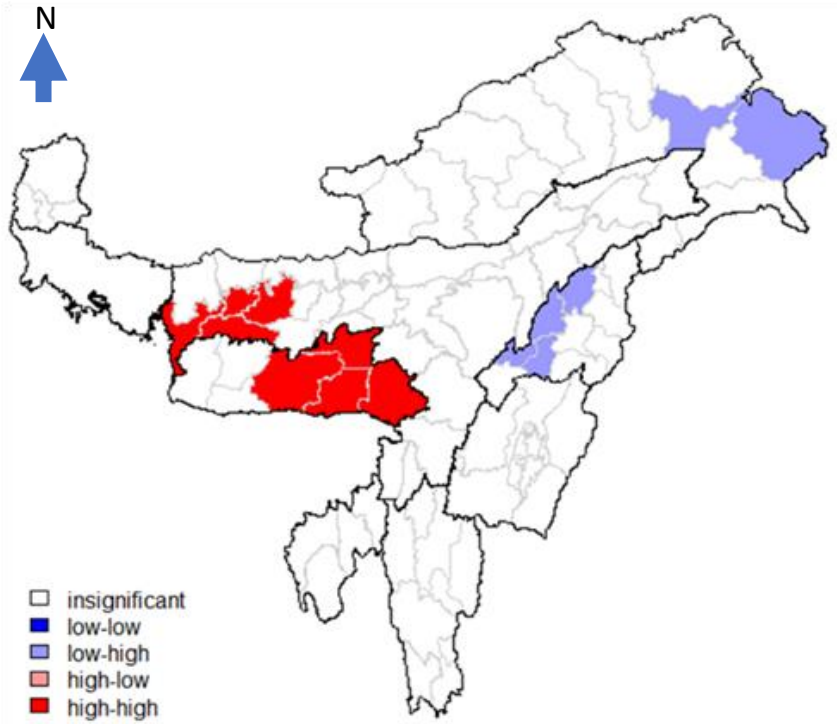
AIMS – to identify statistically significant clusters

Technique – Getis-Ord Gi* (G-I star)

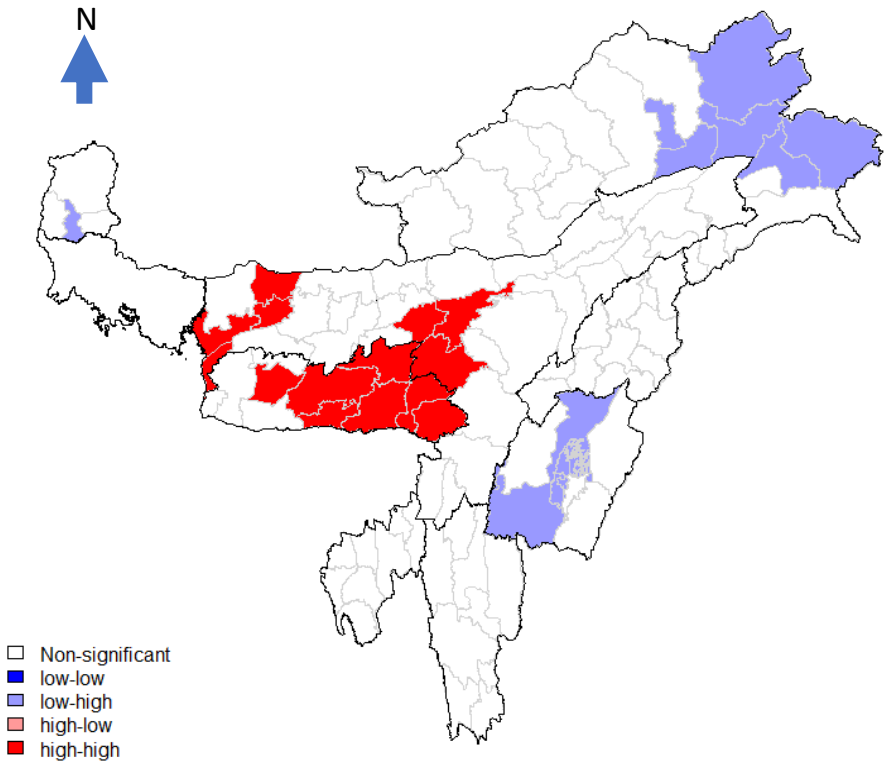


Source: Kh. Jitenkumar Singh et.al.(2025). Identification of spatial hot spots clustering and geographically weighted regression analysis to assess predictors of cesarean section delivery in northeastern states, India. The Journal of Obstetrics and Gynecology of India.

Univariate spatial analysis



a) NFHS-4 (2015-16)



b) NFHS-5 (2019-21)

- This univariate analysis identifying the trends of persistent clusters of stunting in few districts

Source: Kh Jitenkumar Singh et.al.(2022). Identifying the trend of persistent cluster of stunting, wasting, and underweight among children under five years in northeastern states of India. Clinical Epidemiology and Global Health Journal.

A Traditional Regression Model

A)

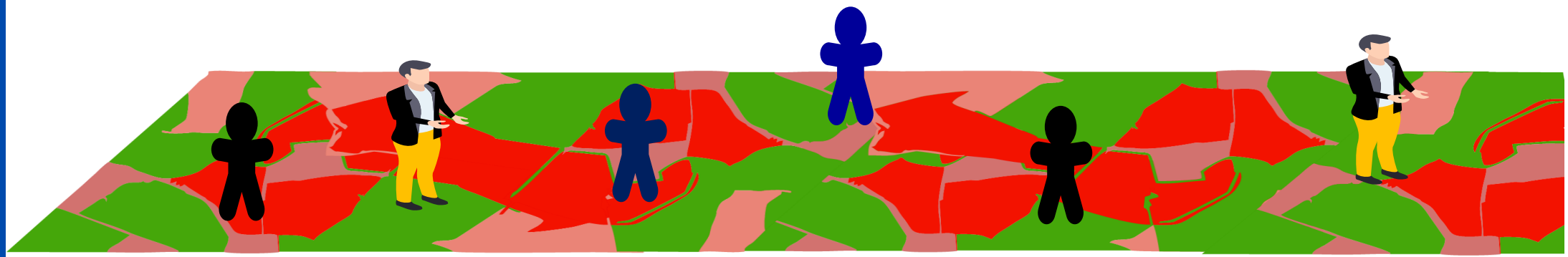
Outcome

Error term

$$y_i = \beta_o + \beta_1 x_{1i} + \ldots + \beta_k x_{ki} + \varepsilon_i$$

Intercept

Vector of covariates



Horizontal

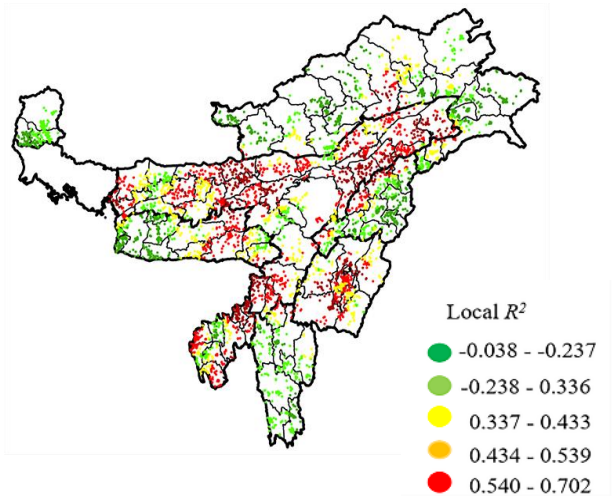
Spatial Effect

B)

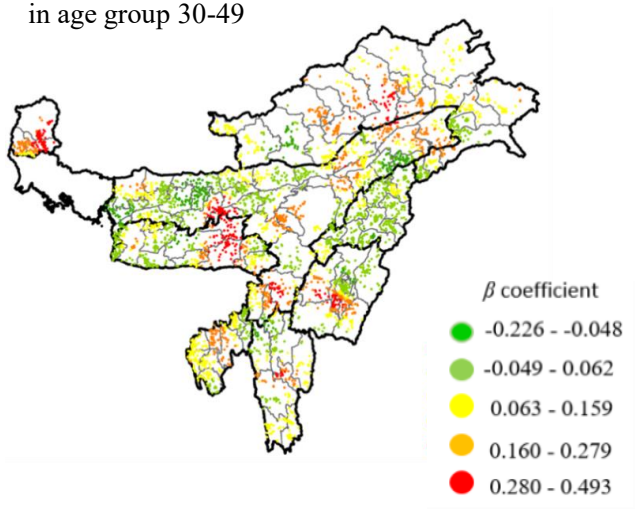
$$\gamma \sim \beta_O + \Sigma\, \beta_i X_i \; + \text{Random effects} \; + \text{Spatial effects}$$

Visualization of R^2 and beta coefficient, Geographically Weighted Regression Analysis (GWR)

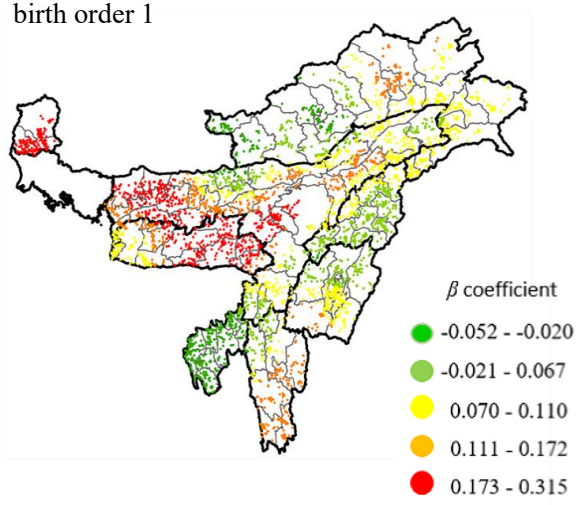
Local R^2 in Northeastern States, India



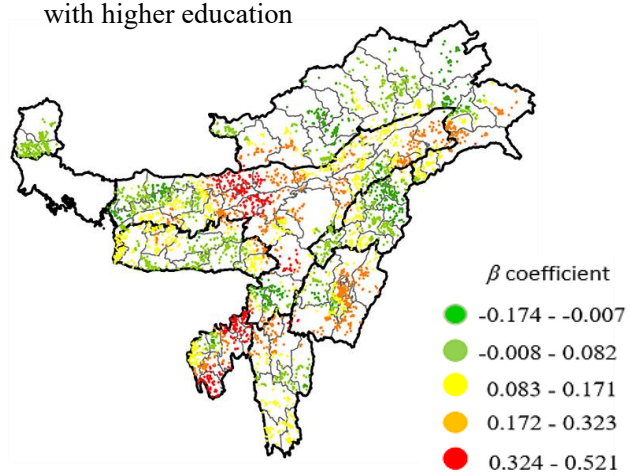
β coefficient of proportion of mothers in age group 30-49



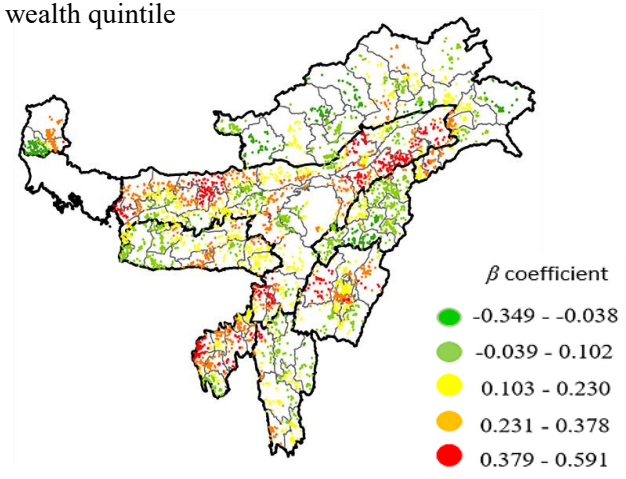
β coefficient of proportion of mothers with child's birth order 1



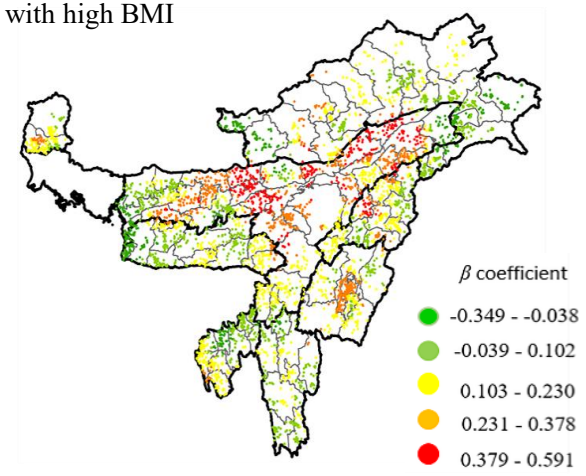
β coefficient of proportion of mothers with higher education



β coefficient of proportion of mothers with highest wealth quintile

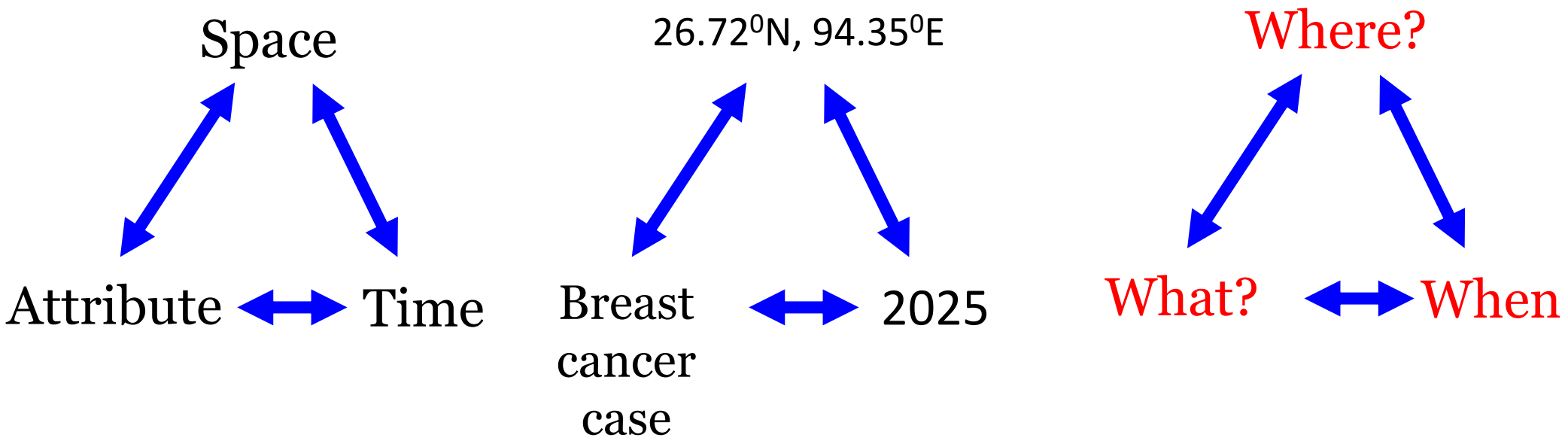


β coefficient of proportion of mothers with high BMI



Source: *Kh. Jitenkumar Singh et.al. (2025):Noncoverage of Hepatitis B immunization among children and its predictors in northeastern states of India: A multiscale geographically weighted regression analysis. Journal of Family Medicine and Primary Care*

Space, Time and Attribute



Inter-relationship between spatial, attribute and time

- In spatial analysis

  focus only spatial variation of the attribute (Space and Attribute)

- In spatio-temporal analysis

  focus spatial variation of the attribute with time (Space, Attribute and Time)

Spatio-temporal - LISA of Caesarean Section delivery in northeastern states (2011 to 2019)

Figure: 2a. LISA of CS delivery, Northeastern States, 2011

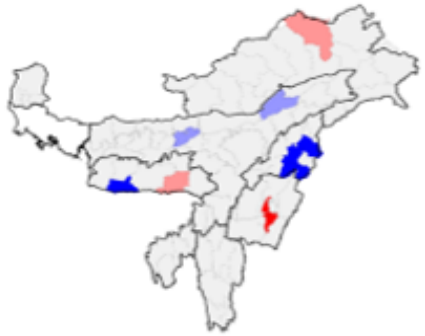


Figure 2b. LISA of CS delivery, Northeastern States, 2012

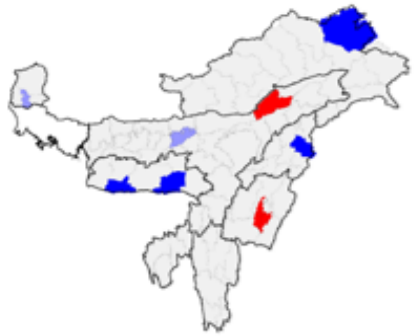


Figure 2c. LISA of CS delivery, Northeastern States, 2013

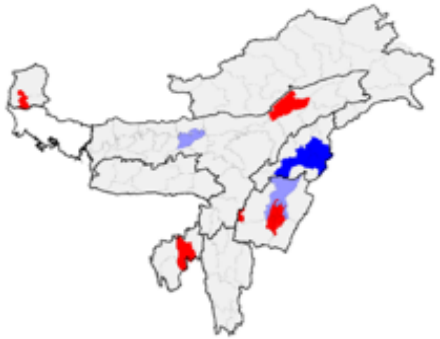


Figure 2d. LISA of CS delivery, Northeastern States, 2014

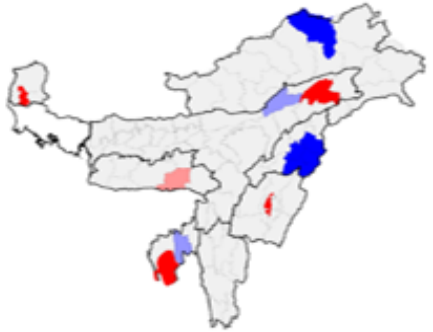


Figure 2e. LISA of CS delivery, Northeastern States, 2015

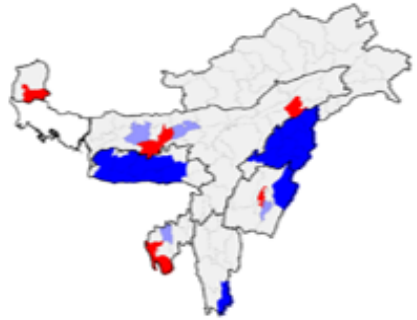


Figure 2f. LISA of CS delivery, Northeastern States, 2016

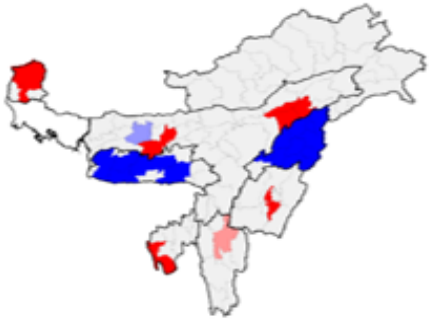


Figure 2g. LISA of CS delivery, Northeastern States, 2017

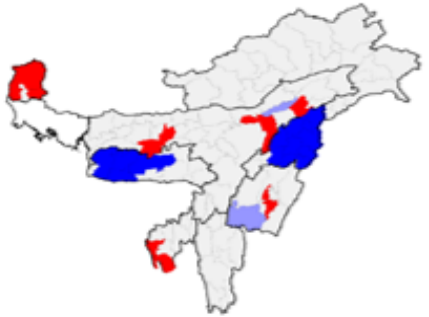


Figure 2h. LISA of CS delivery, Northeastern States, 2018

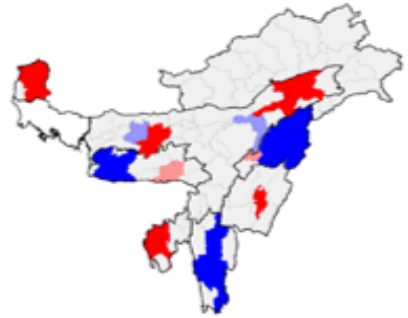


Figure 2i. LISA of CS delivery, Northeastern States, 2019

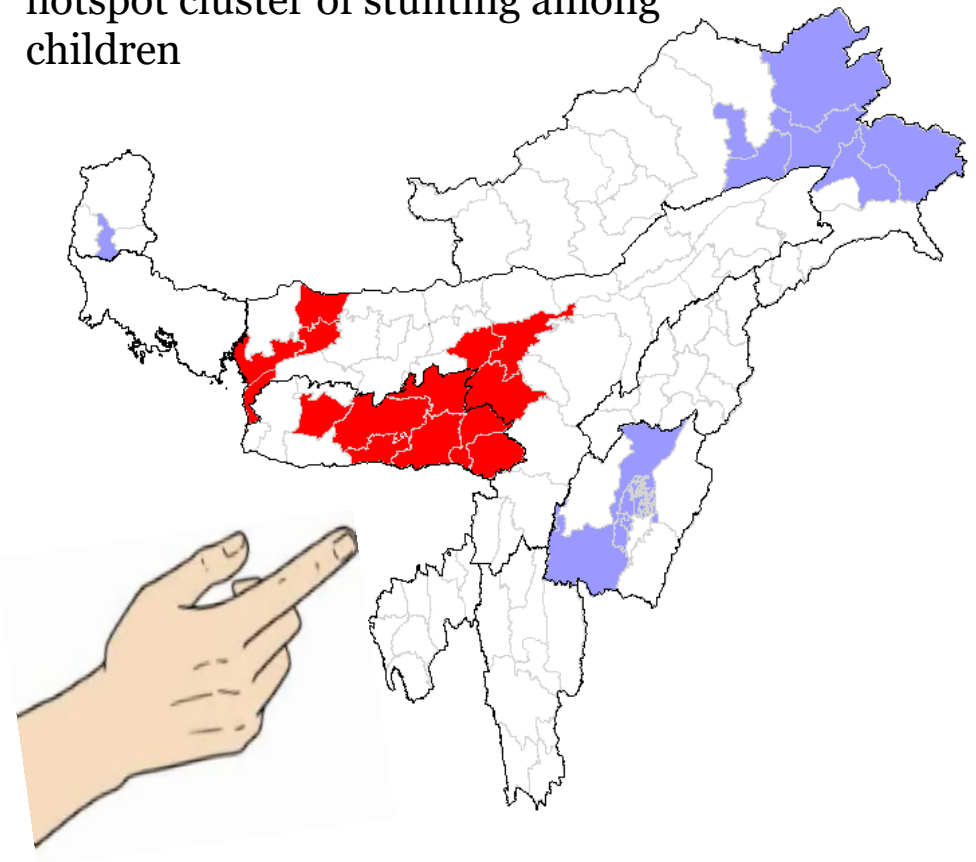
Source: Kh. Jitenkumar Singh, et. al.. (2025). Cesarean section delivery and determining factors in Northeastern State, India: A Bayesian spatio-temporal analysis of areal data. Indian Journal of Medical Research.

Implications for Public Health Policy

Spatial data analysis offers significant advantages for public health:

- *Enables targeted population, screening, and treatment programs.*
- *Guides efficient allocation of limited resources in the program.*
- *Supports monitoring progress towards the elimination of disease targets.*
- *Facilitates understanding of socio-economic and environmental determinants of disease.*

Univariate analysis (LISA) - hotspot cluster of stunting among children



NFHS-5 (2019-21)

Source: Kh Jitenkumar Singh et.al.(2022). Identifying the trend of persistent cluster of stunting, wasting, and underweight among children under five years in northeastern states of India. Clinical Epidemiology and Global Health Journal.

Challenges in Spatial Health Research

Data availability & Quality

Issues like under-reporting, incomplete registries, and missing geographic information.

Causal Inference

Lack of individual-level data can limit the ability to draw causal conclusions.

Capacity Building & Computational demands

Need for increased training and expertise in spatial epidemiology.

Ethical issues with geolocation

Navigating privacy and ethical considerations when mapping sensitive health data.

Opportunities & Future Directions

- Integration with Machine Learning
- Real-time disease surveillance
- Participatory GIS and mHealth (mobile health)
- Climate-Health linkage

Key Takeaways

- Spatial methods reveal hidden geographic patterns of disease distribution and health outcomes.
- Essential for guiding geographically targeted public health interventions, resource allocation and evidence-based public health strategies.
- Must be applied thoughtfully, adhering to principles of epidemiologic rigor.

Thank You