#10/2025

CLUSTERING-BASED APPROACHES FOR ASSESSING SERVICE ACCESS INEQUALITY: A SYSTEMATIC LITERATURE REVIEW OF HOUSEHOLD AND COMMUNITY MICRODATA STUDIES

Main Author: Macapayad, Khryzz Lynn
Co-Author: Salapa, Aristeo

University of Southeastern Philippines, Davao City, 8000
*klmacapayad02202400410@usep.edu.ph, a.salapa@usep.edu.ph

Abstract

Clustering-based methodologies are crucial for studying inequitable access to services in rural contexts. Following the PRISMA guidelines for systematic reviews, we assessed 30 empirical papers out of an initial 253. With a focus on access to basic services, the inclusion criteria were set to the strictest parameters within the boundary of using clustering techniques on granular household or community-level microdata. Of the 83 records that passed the title/abstract screening stage, 30 were retained in the final set after reviewing the full text. The results of the studies demonstrate the ability of clustering techniques to reveal and better describe geographical concentrations in service deficits that are otherwise not visible through analytical means of aggregate data or are often overlooked. Regarding the systematic empirical clustering of service access, this is highly useful for evidence-based planning of public services, particularly in lower- and middle-income countries with a specific geographical focus.

Keywords: clustering methods, service access inequality, household microdata, spatial analysis, CBMS, poverty mapping, geographic targeting, PRISMA systematic review

1. INTRODUCTION

The shortage and ongoing challenges to basic services, including water, sanitation, electricity, healthcare, education, and transportation, pose a significant development hurdle. Dedicated accessibility of such services is crucial, especially in remote rural and island locations and geographies. Goal number 6 of the Sustainable Development Goals (SDGs) of the UN illustrates the commitment to accessibility of basic services to all; however, as the goals also emphasize, inequalities continue to exist. Access (or lack thereof) to basic services is one of the major challenges in areas such as South Africa (David et al., 2018; Shifa & Leibbrandt, 2021) and Ethiopia (Kibret et al., 2024), as well as throughout the greater sub-Saharan Africa region, as numerous studies have demonstrated. The persistent deprivation of basic services in economically disadvantaged areas is, without a doubt, a result of inequalities of all kinds, especially in low- and middle-income countries (LMICs). The most affected are the disadvantaged population (Leibbrandt et al., 2018).

Measuring and understanding the complexities of service deprivation and access inequalities requires a sophisticated approach, one that is devoid of a one-size-fits-all methodology. Geographic information systems (GIS) and geographic spatial analysis, along with a clustering methodology, have proven to be invaluable assets. The great combination of the two methods (i.e., GIS and clustering), such as k-means, hierarchical, and self-organizing mechanisms, will assist in gaining an understanding of the geographic areas in regard to deprivation and also to classify/subset the population on the basis of the access profile on basic services (Wei, Cabrera-Barona, & Blaschke, 2016; Charrahy, Aghamohammadloo, & Shamoradi Samani, 2025). These methods, especially the combination of GIS and advanced clustering, provide the necessary evidence to assist in local area planning and optimize the limited available resources to address spatially defined challenges. Household and community microdata originate from sources of Census-Based Monitoring System (CBMS) surveys, national population censuses, Demographic and Health Surveys (DHS), Living Standards Measurement Surveys (LSMS), and Multiple Indicator Cluster Surveys (MICS) (Shifa & Leibbrandt, 2021; Nguetse Tegoum & Hevi, 2016). The geocoded examples of the datasets configured with clustering techniques enable policymakers and planners to uncover "service deprivation hot spots," identify dendrogram communities with homogeneous deprivation profiles, and allocate resources to socially underserved zones, as well as prioritize equity-oriented development goals through tracer initiatives. Clustering algorithms have been used in India, the Philippines, South Africa, and Bangladesh to map multiple dimensions of poverty and social deprivation, as well as to inform targeted infrastructural programs.

The growing use of clustering algorithms/modules hinges on precise geocoding as the principal foundation for microdata empirical models. A plethora of research works embodies innovative ideas pertaining to geocoded datasets and clustering algorithms; however, a systematic and comprehensive literature review of these eclectic points from varied geographies and methodologies is yet to materialize, which this research intends to achieve. Hence, this research aims to meticulously identify, review, and assess peer-reviewed and non-peer-reviewed literature on the use of sophisticated micropopulation and household community datasets to investigate access inequities across 30 studies. The outcome of the reviewed literature will be the microfoundation on which innovative methods will be manifested, and the empirical evidence will be available for policymakers in the Global South.

2. METHOD

Design. A systematic literature review was conducted to determine the clustering-based approaches used to assess service access inequality using household and community microdata. This study adhered to the criteria of the Preferred Reporting Items for Systematic Reviews and Meta-Analysis (PRISMA 2020), ensuring a rigorous and transparent methodology consistent with international standards for systematic reviews.

Eligibility Criteria. Such studies included those focusing on empirical and quantitative studies using clustering techniques on microdata at the household or community level to analyze inequalities in access to services published in full-text English documents between the years 2010 and 2025. The varied services included in the definition of access were water, sanitation, and electricity; health and education; and transport and ICT services. The geographical focus was global, and particularly on low- and middle-income countries (Sisay, Muche, & Ali, 2020; McCordic, Frayne, Sunu, & Williamson, 2022).

Information Sources. To find suitable and relevant material, this systematic review was performed within the scope of the social sciences, international development, and development studies. To find relevant peer-reviewed journal articles, conference proceedings, and theses, key research databases and search engines, including Scopus and Web of Science as well as Google Scholar, were consulted. We also sought the development institutions, such as the United Nations and the World Bank, to aid our search for unpublished and non-open-access literature. In order to perform systematic and comprehensive searches, we looked for studies written in English and were empirical in the field of social sciences and international development, which also used some form of cluster analysis on microdata from households or communities regarding access to essential services. The review sought to be as comprehensive as the literature would allow. Many adjustments were made in order to improve the ease of retrieval, specificity, and accuracy of the different search engines. The search techniques were customized for the different search interfaces and thesaurus systems of the different platforms.

Table 1. Search space for selected databases.

Source Category	Source Name	Search Method	Date of Search
Online Database	WOS	Abstract, Title, and keywords	2025-11-19
	Scopus	Abstract, Title, and keywords	2025-11-21
Search Engine	Google Scholar	Full Text, Abstract, Title, and keywords	2025-11-23

Search Terms. The search strategy combined three main groups of concepts: clustering methods, basic service access, and household or community microdata. Boolean operators and synonyms were employed to capture diverse terminologies across disciplines while maintaining specificity to service-access-related studies. The first search string was about clustering and service access. It included words like "clustering," "k-means," "unsupervised learning," and "spatial clustering," as well as phrases like "basic services," "service access," and "service deprivation." It also included data source descriptors like "household microdata," "census microdata," "community data," and "CBMS." The second string targeted indicators of basic services, using keywords like "basic service indicators," "access indicators," and "infrastructure indicators," along with "household survey," "community-level data," and "microdata." The third search iteration combined microdata-related terms ("microdata," "household data," "CBMS") with inequality-focused concepts such as "spatial inequality," "infrastructure gaps," and "service inequality." These strings were changed to fit various syntax rules related to field codes and truncation so that the integrity of the concepts and the complete retrieval of relevant literature regarding the assessment of the inequality of service access through the application of clustering methods to microdata would be preserved. This is aligned with the standards of systematic literature reviews as far as search terms are concerned and can be found in academic literature on search strategy formulation.

Study Selection. The study selection process followed a systematic, four-stage protocol, detailed in Table 2, and was conducted in accordance with the PRISMA 2020 guidelines. Figure 1 summarizes the flow and results of this process. A total of 253 records were identified through database searches: Web of Science (n = 112), Scopus (n = 18), and Google Scholar (n = 123). After removing 26 duplicates, 227 unique records were screened by title and abstract. Of these, 140 records were excluded because they did not meet the population, intervention, or study design criteria. The full text of the remaining 87 articles was sought for retrieval. While 13 reports could not be retrieved, the full texts of the 74 successfully retrieved reports were assessed for eligibility by two independent reviewers. During this full-text review, 44 reports were excluded due to specific methodological or topical misalignments (e.g., incorrect outcome, incorrect data type). Consequently, 30 studies met all inclusion criteria and were included in the final synthesis.

Table 2. Stages of the study selection process

Stage	Description
S1	Initial identification through database and search engine searches
S2	Duplicate removal
S3	Title and abstract screening using standardized criteria
S4	Full-text review by one independent reviewer against detailed eligibility criteria

Table 2 illustrates the sequential filtering applied, from initial identification to final inclusion. The most substantial refinement occurred during the full-text review (S4), where the rigorous application of detailed eligibility criteria ensured that the selection was precisely focused on clustering methods applied to microdata for assessing service access inequality. The final distribution of the 30 included articles across the three information sources is presented in Table 3, which shows that Google Scholar contributed the majority of both the initially identified (n = 123) and finally included (n = 16) studies, followed by Web of Science.

Figure 1 (the PRISMA flow diagram) visualizes the complete flow of identification, screening, eligibility, and inclusion. Figure 1 provides a transparent audit trail, depicting the attrition at each decision point—the transition from 87 reports sought to 30 included—and justifying all exclusions with specific reasons, thereby upholding the reproducibility and rigor of the review process.

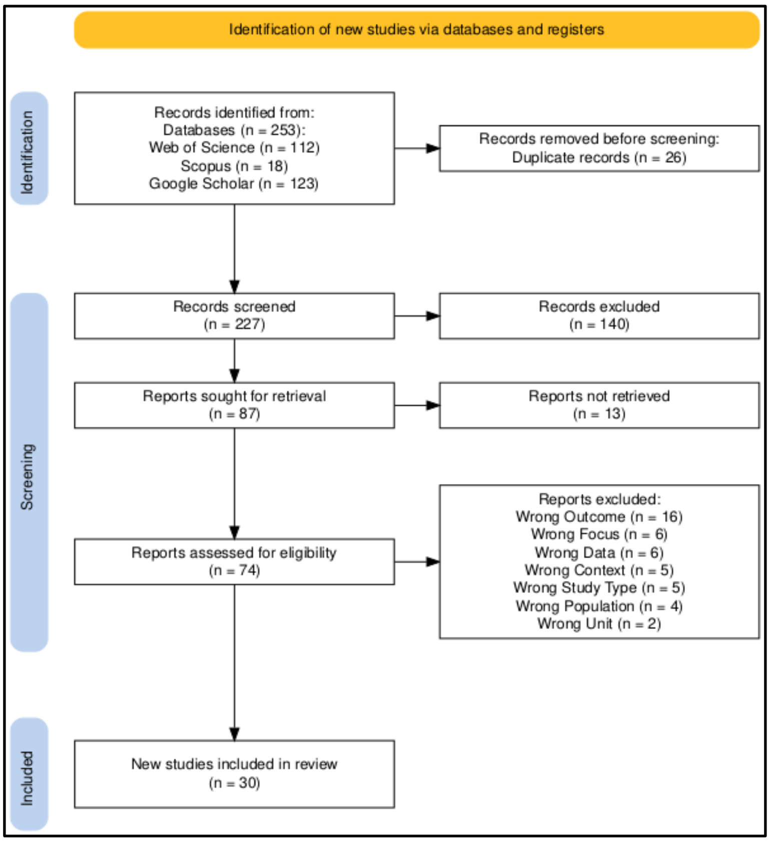

Figure 1. Contextualized PRISMA Flow Diagram of the Study

Table 3 presents the distribution of articles across the three information sources, both before and after the selection process. Table 3 quantifies the contribution of each database, revealing that Google Scholar was the most prolific source, yielding 123 initial records and 16 (53.3%) of the final included studies. This is followed by Web of Science, which contributed 112 initial records, and 11 (36.7%) included studies. Scopus contributed a smaller share. The initial dataset consisted of 18 records, of which 3 (10%) were included in the final studies. This distribution highlights the complementary role of specialized databases (Web of Science, Scopus) and broad search engines (Google Scholar) in achieving a comprehensive literature search.

Table 3. Distribution of Articles Before and After the Selection Process

Source	Before	After
Web of Science (WOS)	112	11
Scopus	18	3
Google Scholar	123	16
Total	**253**	**30**

Data extraction and synthesis. The process of extracting and synthesizing the collected data was performed by carefully reading all 30 articles, and MS Excel was used to record the extracted data and results. The summary of SLR was presented in column headlining themes, article titles, authors, years of publication, research design, participants/respondents or related sources, observed variables, indicators of basic service access, data used, a brief description of the study, and the findings of the study in relation to the clustering approaches for analyzing inequalities in access to basic services.

3. RESULTS AND DISCUSSION

3.1 Use of Household/Community Microdata in Clustering Analyses for Service-Access Inequality

Aggregated Multidimensional Deprivation Clustering. This illustrates the initial idea, in which microdata on service access and assets, education, and health, among others, are geographically aggregated at the district, municipality, or ward level, and composite deprivation scores are created. These scores are then used by clustering algorithms to prioritize the allocation of policy resources to administrative units with the same level of deprivation. For instance, Fadilah et al. (2025) aggregated household surveys from the Indonesian Central Bureau of Statistics to the district level regarding access to sanitation and clean water and then performed a k-means clustering of the districts in Papua. Along similar lines, Katumba (2018), with the aid of SPA (the South African Poverty Multidimensional Index), a composite of census microdata containing access to water, sanitation, and energy, conducted a spatial statistical analysis of the deprivation concentration of Gauteng. For instance, in a Cameroon Multidimensional Poverty Study, what Nguetse Tegoum and Hevi (2016) did was use the Alkire-Foster method on microdata household surveys, which clustered households and regions vertically and horizontally, and then revealed the overlapping deprivations in health, education, and other basic service deprivations (i.e., water and sanitation). This method combines thousands of individual household records to create a streamlined composite of the most deprived areas, thereby enhancing the effectiveness of region prioritization and resource allocation.

Geospatial Hotspot and Autocorrelation Analysis. Because of the geographic identifiers (GPS coordinates, administrative codes, etc.) embedded in microdata, this accurately pinpoints and statistically maps the concentration (hot and cold spots) and level of service accessibility. This method accurately depicts the geography of inequality. Access to health services is the most widely researched in this area. Sisay et al. (2020) examined the geospatial distribution of institutional delivery using georeferenced Ethiopian DHS data and found significant spatial autocorrelation in the low-access cluster concentration. The finding was built upon by Kibret et al. (2024), who used the same DHS data to construct the geography of maternal and neonatal healthcare and, along with cluster analysis of the healthcare, added variables related to women's empowerment. Deitz and Meehan (2019) used the American Community Survey microdata in a high-income setting to map the geography of poverty (i.e., the lack of complete plumbing facilities) and plumbing poverty, illustrating inequality in the landscape along ethnic and geographic lines. This method, similar to the others, is based on microdata and provides information that confirms the geography of opportunity and the geography of deprivation as major mechanisms of inequality; furthermore, the microdata serves as the precise information needed for meaningful spatial analysis, as noted by Shifa and Leibbrandt (2024) in their review.

Household-Level Vulnerability Profile Clustering. It utilizes microdata to describe how the interaction of different household characteristics, which include service availability, asset ownership, exposure to shocks, and socio-demographic characteristics, defines multiple layers of synergistic vulnerability. At this point, microdata are not for aggregation but for direct clustering of households of interest and their related characteristics. McCordic et al. (2022) used original household surveys collected from five cities in the Global South to assess how the simultaneous lack of water, electricity, and sanitation, as well as the clustering of these household amenities, aggravates food insecurity and demonstrates different patterns of vulnerability. Similarly, resilience frameworks utilize microdata to cluster households based on their ability to withstand shocks. Olawuyi and Mushunje (2023) used the RIMA tool of the FAO to household survey data from South Africa (where the "Access to Basic Services" pillar is one of the main components used to cluster The study aims to gauge the resilience of households. Muriuki et al. (2024) used household panel data from Uganda and Malawi to identify different groups of households with different strategies and difficulties for resilience and discovered that low resilience capacity (which was found to be poor service access) clusters in a systematic way. This interpretation of the data uses households as the unit of analysis and employs clustering to describe groups of households defined by their complex synergistic relationships, rather than by a single deprivation.

Ultimately, community and household microdata are crucial for the analysis of service-access inequities. They contain the relevant attributes for Forming Aggregated Multidimensional Deprivation Clustering, the geocoded data needed for Hotspot Analysis and Spatial Autocorrelation, and the integrated records necessary to construct Clusters of Vulnerability Profile Households. The analysis aims to identify deprived regions, spatial injustice, and complex risk profiles, which are examined using microdata at both the unit and higher levels. They offer a precise and evidence-based understanding of the stratification and concentration of inequities in basic service access across various populations and geographies.

3.2 Indicators of Basic Service Access in Household- and Community-Level Microdata Studies

Water, Sanitation, and Hygiene (WASH) Indicators. One case in point is drinking water, which is usually classified in terms of safety and distance. Some research makes a difference between 'improved' sources (piped water, protected wells) and 'unimproved' sources (unprotected wells, surface water) (Fadilah et al, 2025; Nguetse Tegoum & Hevi, 2016). One of the better measures is what McCordic et al. (2022) refer to as having 'reliable access to piped water.' This entails service continuity, rather than the mere type of source. As pertains to sanitation, the main measure is the type of toilet facility that is available, which is classified as improved (flush/pour-flush to sewer/septic tank or ventilated improved pit latrines) and shared or unimproved (Fadilah et al, 2025; Katumba, 2018). In high-income countries that include the U.S., Deitz and Meehan (2019) apply a specific indicator to measure one's plumbing poverty status, which is having complete plumbing facilities along with hot and cold piped water, a bathtub or shower, and a flush toilet.

Energy and Electrification Indicators. Electrical access is one of the most common indicators of multidimensional poverty (Fadilah et al., 2025; Katumba, 2018; Parreno Jr. and Del Mundo, 2015). However, some studies have gone beyond mere access and assessed reliability. While some studies prioritize the number and length of outages, McCordic et al. (2022) determine the others to be essential to urban vulnerability. A common indicator of a household's access to electricity is the primary source of lighting used, be it electricity, kerosene, candles, or batteries. This type of indicator is also common in surveys (Nguetse Tegoum & Hevi, 2016).

Healthcare Access and Utilization Indicators. Geospatial studies often analyze physical access in terms of distance or time required to travel to the closest healthcare provider (Popoola et al., 2016; Sisay et al., 2020). Self-reporting from surveys serves as the basis for the indicators of utilization, such as whether a birth was attended by a qualified health worker or whether the delivery occurred in a health institution (Kibret et al., 2024; Sisay et al., 2020; Wabiri et al., 2016). Vaccine acquisition, as researched by Saha et al. (2018) for cholera, is a critical preventive healthcare measure. Additionally, the health resource density (e.g., the number of psychiatrists per population, as used by Niranjan et al. (2025)) is a community-level measure of access on the supply side.

Digital and Informational Service Indicators. Muriuki et al. and Olawuyi and Mushunje (2023) measure outreach to mobile and internet technologies (ICT) as a key indicator. Access to markets, information, and social safety nets (and subsequently improving livelihoods) has become a frontier of mobile phones and internet access, and Muriuki and Olawuyi (2023) empirically measure e-governance awareness, a measure that intersects digital access and civic engagement. Zhang and Fan (2025) focus on location and spatial concentration measurements of healthcare services, among others, to measure specialized service equity on public electric vehicle (EV) charging infrastructure.

Composite Infrastructure and Resilience Indicators. This indicator aggregates vulnerabilities or deprivation related to the composite access dimensions. Studies often index WASH, energy, and housing. For example, Balakrishnan et al. (n.d.) assess the availability of synergistic regional infrastructure. The South African Multidimensional Poverty Index (SAMPI) was used by Katumba (2018) and incorporates assets and services such as access to water, sanitation, and energy. Additionally, resilience frameworks, such as the RIMA tool used by Olawuyi and

Mushunje (2023), include access to core basic services as a "pillar" of household resilience capacity and measure it using a specific indicator basket of the above. McCordic et al. (2020) explicitly model bundles of access to infrastructure (water, electricity, and sanitation) and their constraining effects on food security, thereby increasing vulnerability to it.

Finally, microdata studies have a range of indicators to analyze service-access inequities. These indicators encompass direct and physical metrics, such as the source of drinking water and access to electricity, as well as behavioral measures of institutional delivery and vaccination uptake, and even modern indicators like mobile phone ownership. The clustered or composite indices alone or in bundles dictate the identification of the disadvantaged cluster and the systemic understanding of how depriving factors interlace to create inequality.

3.3 Clustering Approaches for Analyzing Inequalities in Access to Basic Services

Geographic and Spatial Clustering. This theme encompasses methods that utilize geography to some extent in determining areas of extreme poverty or inadequate services. The main premise is that inequalities occur in the same locations, meaning that both the inequalities and the areas experiencing deprivation are clustered together. One of the most basic methods of study is the spatial autocorrelation measurement, also known as the LISA, which is used to determine statistically significant areas where poverty disparities are present (hot spots) and areas of wealth (cold spots). For instance, David et al. (2018) utilized municipality-level spatial autocorrelation to demonstrate the high concentration of poverty and inequality in South Africa, as they showed that deprived municipalities were often neighbors to each other. Exactly the same way, Deitz and Meehan (2019) used the hot spot analysis (Getis-Ord Gi*) to demonstrate the extremely high level of "plumbing poverty" in the United States, where there are significant and profound racial and geographical inequalities in the households. The same concept of clustering poverty is also used to demonstrate access (or lack thereof) to healthcare services. Niranjan et al. (2025) used deprivation geospatial analysis to inequitable distribution of mental health facilities and psychiatrists in the center of India, while Kibret et al. (2024) and Sisay et al. (2020) employed the spatial clustering phenomenon to describe the geographical distribution of the continuity of maternal and neonatal health services in Ethiopia. The scale of analysis is often overlooked, as Wei et al. (2016) demonstrated that the degree of public service inequality experienced varies by the neighborhood scale chosen as the unit of analysis.

Socioeconomic and Multidimensional Clustering. This theme pertains to the usage of socioeconomic information to create typologies of deprivation beyond income. Socioeconomic data across geospatial units can be aggregated to create typologies of deprivation. Well-being consists of a myriad of dimensions: education, health, and living conditions. Most often, the K-means algorithm is the most common technique used to randomly designate a portion of a country's districts to separate clusters. Fadilah et al. (2025) used K-means clustering to classify Papua's districts along a number of poverty indicators so that they could identify and visualize the districts exhibiting a similar deprivation profile. For example, Oliveira (2021) used K-means and other unsupervised learning techniques on open geospatial datasets to try to classify different typologies of impoverished areas without focusing on the quantitative aspects, which is similar to the Multidimensional Poverty Index (MPI) that employs a number of different statistical techniques to assess the level of concentration of multiple deprivations. One example is Katumba (2018), who used statistical techniques to assess the extent and concentration of what is called multidimensional poverty in Gauteng, South Africa. The applications of clustering in multidimensional poverty frameworks are apparent in Nguetse Tegoum and Hevi (2016) and Jiang et al. (2025), focusing on child poverty and assessing poverty mitigation strategies. This approach assesses the various types of deprivations of the households and areas/regions of study using multiple indicators.

Infrastructure and Service-Specific Clustering. This theme focuses on conducting tailored analyses of accessibility inequalities related to basic services or infrastructure systems for the purpose of clustering. The service in question, along with its accessibility and spatial variables, drives the clustering. One major example is food security, where clusters are defined by access to infrastructure and vulnerability. In the cities of the global South, McCordic, Frayne, and Sgro (2020) and McCordic et al. (2022) analyzed the spatial clustering of disrupted service access and the concomitant household vulnerability to food insecurity in urban areas, particularly in relation to the basic services of water and electricity. In the field of infrastructure planning, clustering algorithms are used for optimization. In a planning support tool for rural electrification, Parreno Jr. and Del Mundo (2015) employed a clustering algorithm to iteratively develop and optimize household groupings for the spatially balanced rollout of electrification. Similarly, Gong et al. (2022) identified and characterized underserved clusters, which described spatial imbalances in the distribution of services to the elderly, indicating inequitable service provision and care. Other service-specific analyses include Popoola et al. (2016) on the distribution of health facilities in Nigeria, Zhang and Fan (2025) on the inequity of charging service infrastructure for electric vehicles, and Balakrishnan et al. (n.d.) on the synergistic influence of regional infrastructure on community exposures. The clustering in these analyses highlights the specific service inequities for targeted investments.

To summarize, the classifications are interdisciplinary in nature because of the use of basic services, spatial econometrics, and cluster machine learning. Spatial and Geographic Clustering (David et al., 2018; Deitz & Meehan, 2019) provides the primary location of inequality and the "where" it comes from. Socioeconomic and Multidimensional Clustering (Fadilah et al., 2025; 2018), Katumba responds to the "who and how" of deprivation by developing new typologies. Last, what is missing? Infrastructure and Service-Specific Clustering (McCordic et al., 2020; Parreno Jr. & Del Mundo, 2015) addresses the "what is absent" by specializing in particular delivery service systems. Altogether, these thematic strands provide diagnostic insights into the inequality landscape and shift policy from a universal to a targeted approach.

4. CONCLUSION

This systematic literature review elucidates how household and community microdata are leveraged within three distinct, yet often complementary, analytical paradigms to diagnose and characterize service-access inequality. The thematic synthesis reveals that clustering is not a single method; instead, it is employed strategically, based on the research question, the data structure, and the policy goal. The primary themes identified—Aggregated Multidimensional Deprivation Clustering, Geospatial Hotspot and Autocorrelation Analysis, and Household-Level Vulnerability Profile Clustering—each address a different facet of the inequality puzzle, from mapping regional deprivation typologies to pinpointing precise spatial gaps and modeling synergistic household risks. These approaches collectively transform granular microdata into actionable intelligence, enabling the shift from broad-brush policies to geographically and demographically targeted interventions. To provide an in-depth examination of the empirical evidence supporting this synthesis, Table 4 presents the 30 reviewed studies. It details each study's thematic focus, methodology, specific data sources, the service-access indicators used, and its key findings on inequality and clustering, offering a consolidated reference for the evidence base discussed throughout this review.

Table 4. Comprehensive Summary of Reviewed Studies on Clustering Analyses for Service-Access Inequality

Theme	Article Title	Author(s)	Year	Research Design	Participants/Data Sources	Variables Observed	Indicators of Basic Service Access	Data Used	Brief Description	Key Finding on Inequality/Clustering
Geographic & Spatial Clustering	1. Spatial poverty and inequality in South Africa: A municipality level analysis	David, A., et al.	2018	Quantitative, Spatial Econometrics	Municipality-level census & survey data (Stats SA)	Income, expenditure, access to services, education.	Access to services (type not specified in title, but implied from census), Education attainment.	South African Census data; National survey microdata aggregated to municipality level.	Analyzes spatial distribution and clustering of poverty and inequality across South African municipalities.	Found strong spatial autocorrelation; poor municipalities cluster together in specific regions (e.g., Eastern Cape), while wealthy ones cluster in metros.
	2. Plumbing poverty: mapping hot spots of racial and geographic inequality in US household water insecurity	Deitz, S., & Meehan, K.	2019	Geospatial Analysis	US Census ACS & housing data	Plumbing completeness, poverty, race, geography.	Complete plumbing facilities (hot/cold piped water, bathtub/shower, flush toilet).	American Community Survey (ACS) microdata (household level).	Maps clusters of incomplete plumbing and water poverty in the United States.	Identified significant hot spots of "plumbing poverty" in specific regions (e.g., Alaska, W. Navajo Nation) and among racial minorities, revealing stark spatial and racial inequality.
	3. Mapping of geographic inequality in mental health care... in Indore division	Niranjan, V., et al.	2025	Geospatial Analysis	Govt. health facility data, population census	Location of psychiatrists & mental health facilities, population.	Physical proximity/availability of mental health facilities and specialists.	Government administrative data on health facility and psychiatrist locations; Census population data.	Maps and analyzes the spatial equity of mental healthcare resources across seven districts.	Found high geographic inequality, with facilities and specialists heavily clustered in the core urban district, creating large underserved peripheral clusters.

Theme	Article Title	Author(s)	Year	Research Design	Participants/Data Sources	Variables Observed	Indicators of Basic Service Access	Data Used	Brief Description	Key Finding on Inequality/Clustering
	4. Spatial patterns of maternal and neonatal continuum of care use...	Kibret, G.D., et al.	2024	Spatial Statistical Analysis	Ethiopian Demographic & Health Survey	Maternal healthcare service utilization, women's empowerment.	Utilization of maternal & neonatal continuum of care (4+ ANC visits, skilled birth attendance, postnatal check).	Ethiopian Demographic and Health Survey (EDHS) microdata (individual women).	Examines spatial clustering of continuum of care completion and its correlation with empowerment.	Identified significant spatial clusters (hot/cold spots) of service use; low-use clusters correlated with lower women's empowerment indicators.
	5. Spatial distribution... of institutional delivery among... women in Ethiopia	Sisay, D., et al.	2020	Spatial Statistical Analysis	Ethiopian Demographic & Health Survey	Institutional delivery service use.	Institutional delivery (childbirth in a health facility).	Ethiopian Demographic and Health Survey (EDHS) microdata (individual women).	Analyzes the spatial pattern and factors associated with institutional childbirth.	Found non-random spatial distribution with significant high-rate and low-rate clusters, highlighting regional inequalities.
	6. Local geographic variation of public services inequality: does the neighborhood scale matter?	Wei, C., et al.	2016	Multi-scale Geospatial Analysis	Urban infrastructure & census data (Vienna, Austria)	Access to public services (schools, doctors, transit).	Physical access/proximity to schools, doctors, and public transit stops.	Urban infrastructure GIS data; Census data for population denominators.	Tests how the measured level of inequality in service access changes with the scale of analysis.	Found that inequality measures are highly sensitive to neighborhood scale, emphasizing the need for scale-aware planning.

Theme	Article Title	Author(s)	Year	Research Design	Participants/Data Sources	Variables Observed	Indicators of Basic Service Access	Data Used	Brief Description	Key Finding on Inequality/Clustering
	19. Spatial Modeling and Analysis of Poverty with a Focus on Spatial Inequalities	Charrahy, Z., et al.	2025	Spatial Modeling (GIS)	Socioeconomic & spatial data (Zanjan Province)	Poverty indices, geographic (slope, elevation, distance to city).	Implied access via poverty indices; Distance to urban centers as proxy for service availability.	Provincial socioeconomic survey/census microdata; GIS layers (slope, elevation, roads).	Integrates socioeconomic and environmental factors in a spatial model to analyze poverty.	Poverty clusters were strongly associated with less-favourable geographic conditions (e.g., higher elevation, distance from urban centers).
	20. Using spatial, hierarchical, and econometric models... to examine food security.	Lopéz-Carr, A.C., et al.	2017	Mixed-Methods, Spatial Modeling	Survey & spatial data (Tanzania)	Food security, market access, land use.	Market access (distance/travel time) as a proxy for access to goods and services.	Primary household survey microdata; GIS data on road networks and market locations.	Combines spatial and econometric models to understand food security determinants in data-poor areas.	Found food insecurity clusters in areas with poor market access, identifiable through spatial modeling.
	26. Socioeconomic drivers of vaccine uptake...	Saha, A., et al.	2018	Secondary Analysis (Cluster RCT)	Data from a cholera vaccine trial (Bangladesh)	Vaccine uptake, wealth, education, distance to site.	Vaccine uptake (cholera); Distance to vaccination site.	Microdata from a cluster randomized controlled trial (household/individual level).	Analyzes how socioeconomic factors influenced uptake within a geographic cluster trial.	Found uptake was not random but clustered by socioeconomic status within geographic clusters, revealing micro-inequalities.

Theme	Article Title	Author(s)	Year	Research Design	Participants/Data Sources	Variables Observed	Indicators of Basic Service Access	Data Used	Brief Description	Key Finding on Inequality/Clustering
	27. Spatial inequality in sub-Saharan Africa.	Shifa, M., & Leibbrandt, M.	2024	Review & Macro-Analysis	Literature and spatial data review	Spatial inequality trends.	Various (review of literature on economic activity, service access).	Synthesis of existing studies and spatial datasets.	Reviews the literature and evidence on spatial inequality across SSA.	Confirms that spatial clustering of economic activity and opportunity is a primary driver of continental inequality.
	30. Spatial Patterns and Developmental Determinants of Women's and Children's Nutrition in India	Waghmare, H., et al.	2025	Geospatial & Statistical Analysis	National Family Health Survey (NFHS)	Nutritional outcomes (stunting, wasting, anemia).	Nutritional outcomes as indirect indicators of access to health, food, and WASH services.	National Family Health Survey (NFHS) microdata (household, women, children).	Maps and analyzes spatial patterns of malnutrition and its determinants.	Identified strong spatial clustering (hot spots) of poor nutritional outcomes, closely linked to regional developmental disparities.
Infrastructure & Service-Specific	13. The household food security implications of disrupted access to basic services...	McCordic, C., et al.	2022	Quantitative, Regression Analysis	Household surveys in 5 Global South cities	Food security, access to water, electricity, sanitation.	Reliable access to piped water, electricity grid, improved sanitation; Frequency of service disruptions.	Original household survey microdata collected in five cities.	Examines how disruptions in basic service access correlate with household food insecurity.	Found that disruptions in water and electricity access cluster spatially and significantly increase odds of food insecurity.

Theme	Article Title	Author(s)	Year	Research Design	Participants/Data Sources	Variables Observed	Indicators of Basic Service Access	Data Used	Brief Description	Key Finding on Inequality/Clustering
	14. The Role of Infrastructure Access in Urban Household Vulnerability to Food Insecurity...	McCordic, C., et al.	2020	Quantitative, Vulnerability Index	Household surveys in Southern African cities	Food insecurity, asset ownership, infrastructure access.	Access to water infrastructure, electricity, sanitation.	Original household survey microdata from Southern African cities.	Develops a vulnerability index linking infrastructure access to food security risk.	Households with poor access to water and energy infrastructure cluster into high-vulnerability groups.
	15. Clustering algorithm as A planning support tool for rural electrification optimization	Parreno Jr, R.P., & Del Mundo, R.	2015	Applied Algorithm Design	Simulated/GIS data of households	Household location, projected demand.	Access to electricity grid (binary: connected/not connected).	Simulated household point location data; GIS data.	Proposes a clustering algorithm to group households for optimal off-grid electrification planning.	Demonstrated that clustering households can minimize infrastructure costs and prioritize service rollout to the most viable clusters first.
	16. Quantifying the Imbalance of Spatial Distribution of Elderly Service with Multi-source Data	Gong, P., et al.	2022	Geospatial Analysis	POI, population, road network data	Location of elderly services, elderly population density.	Physical proximity/availability of elderly care service facilities.	Points of Interest (POI) data for services; Census data for elderly population; Road network data.	Models and quantifies the spatial mismatch between elderly population and service facilities.	Revealed significant service deserts (clusters of high elderly population with low service access) versus service-rich clusters.

Theme	Article Title	Author(s)	Year	Research Design	Participants/Data Sources	Variables Observed	Indicators of Basic Service Access	Data Used	Brief Description	Key Finding on Inequality/Clustering
	17. Equity evaluation of community-based public EV charging services...	Zhang, Y., & Fan, Y.	2025	Geospatial & Equity Analysis	EV charging station data, census, land use (Sacramento)	Location of EV chargers, demographic variables (income, race).	Physical proximity/availability of public Electric Vehicle (EV) charging stations.	EV charging station location data; American Community Survey (ACS) census tract data.	Evaluates the spatial equity of public EV charging infrastructure distribution.	Found charging stations are inequitably distributed, clustering in higher-income and majority-white census tracts.
	18. Empirical Evidence for Synergistic Influence of Regional Infrastructure Availability and Access...	Balakrishnan, S., et al.	n.d.	Quantitative, Econometric	Regional data	Infrastructure indices, community vulnerability metrics.	Composite indices of transport, communication, energy, and social infrastructure access.	Aggregated regional-level data from official statistics/surveys.	Analyzes how the combination of different infrastructure types influences community vulnerability.	Suggests vulnerability clusters in regions where multiple infrastructure deficits (e.g., transport + communication) coincide synergistically.
	22. Community Resilience Capacity, Environmental Factors and Performance of Food Security Projects...	Ekiru, M.N.	2020	Mixed-Methods, Case Study	Household surveys, project data (Turkana, Kenya)	Resilience capacity, environmental shocks, food security.	Implied access via asset ownership and community infrastructure.	Primary household survey microdata; project monitoring data.	Examines how community resilience and environmental factors affect food project success.	Found that low resilience capacity (tied to infrastructure and asset access) clusters in certain communities, predicting project performance.

Theme	Article Title	Author(s)	Year	Research Design	Participants/Data Sources	Variables Observed	Indicators of Basic Service Access	Data Used	Brief Description	Key Finding on Inequality/Clustering
	28. Predicting climate smart agriculture (CSA) practices using machine learning...	Noma, F., & Babu, S.	2024	Predictive ML	Farmer survey data	Adoption of CSA practices, access to info, credit, extension.	Access to agricultural extension services and information.	Primary survey microdata of farmers.	Uses ML to predict adoption of climate-smart agricultural practices.	Found access to information and extension services are key clustering variables; farmers without access form a distinct non-adopter cluster.
	29. Geospatial Analysis of the Distribution of Health Facilities in Peri-urban area...	Popoola, O., et al.	2016	Geospatial Analysis	Health facility locations, population data	Location of health facilities, population distribution.	Physical proximity/availability of health facilities.	Health facility location data; Population census data.	Maps and analyzes the spatial distribution of health facilities relative to population.	Identified service gaps and clusters of high population density with low facility access, indicating locational inequality.
Socioeconomic & Multidimensional Clustering	7. APPLYING K-MEANS CLUSTERING FOR GROUPING PAPUA'S DISTRICTS BASED ON POVERTY INDICATORS	Fadilah, Y.C., et al.	2025	Quantitative, Unsupervised ML (K-means)	District-level poverty data (BPS Papua)	Poverty depth, severity, expenditure, school participation.	Access to improved sanitation, improved drinking water, electricity (as components of poverty index).	Aggregated district-level data from the Central Bureau of Statistics (BPS), derived from survey/census microdata.	Uses K-means algorithm to cluster districts into distinct groups based on multiple poverty indicators.	Clustered districts into 3 groups (High, Medium, Low poverty severity), revealing clear geographic patterns of multidimensional deprivation.

Theme	Article Title	Author(s)	Year	Research Design	Participants/Data Sources	Variables Observed	Indicators of Basic Service Access	Data Used	Brief Description	Key Finding on Inequality/Clustering
	8. The Diversity of Deprived Areas: Applications of Unsupervised Machine Learning...	Oliveira, L.T.	2021	Quantitative, Unsupervised ML	Open geodata (OpenStreetMap, census)	Building density, land use, road types, socioeconomic data.	Proxy indicators: building quality, road network type, proximity to amenities.	OpenStreetMap (OSM) geodata; Aggregated census/survey data.	Applies unsupervised learning to identify and characterize different typologies of deprived urban areas.	Found that deprived areas are heterogeneous; clustering revealed distinct types (e.g., dense inner-city vs. peripheral informal), requiring tailored policies.
	9. Spatial statistical analyses to assess... multidimensional poverty in Gauteng...	Katumba, S.	2018	Spatial Statistical Analysis	Census data, South African MPI	Health, education, living standards (MPI indicators).	Access to water, sanitation, electricity, refuse removal (as MPI living standards indicators).	South African Census microdata aggregated to small area level; South African Multidimensional Poverty Index (SAMPI).	Uses spatial stats to analyze the extent and concentration of multidimensional poverty.	Identified significant spatial clustering of multidimensional poverty, with hot spots in townships and informal settlements.
	10. Child poverty and household poverty in Cameroon: a multidimensional approach	Nguetse Tegoum, P., & Hevi, K.D.	2016	Quantitative, Multidimensional Analysis	Cameroon Household Survey	Health, education, nutrition, housing, water, sanitation.	Access to improved water, improved sanitation, electricity.	Cameroon Household Survey (CHS) microdata.	Measures child and household poverty using Alkire-Foster method, analyzing	Found different clustering of deprivations; child poverty was higher and driven by different factors

Theme	Article Title	Author(s)	Year	Research Design	Participants/Data Sources	Variables Observed	Indicators of Basic Service Access	Data Used	Brief Description	Key Finding on Inequality/Clustering
									deprivation profiles.	than general household poverty.
	11. Deeper Effects of fiscal multidimensional poverty reduction...	Jiang, W., et al.	2025	Quantitative, Econometric Analysis	Household survey & fiscal data	Multidimensional poverty, fiscal transfers, household characteristics.	Components of a multidimensional poverty index (including service access).	Household survey microdata linked with fiscal program data.	Analyzes how household characteristics influence the effectiveness of fiscal poverty reduction.	Found elite capture and financial lags create unequal impacts, effectively clustering benefits away from the most deprived households.
	12. Area disparity of income and expenditure in Sri Lanka...	Baba, Y., et al.	2015	Quantitative, Statistical Analysis	Household Income & Expenditure Survey microdata	Income, expenditure patterns.	Expenditure on basic services (utilities, transport, communication).	Household Income and Expenditure Survey (HIES) microdata.	Analyzes regional disparities in economic well-being using detailed household data.	Identified distinct regional clusters with significantly different income and consumption patterns, indicating economic inequality.

Theme	Article Title	Author(s)	Year	Research Design	Participants/Data Sources	Variables Observed	Indicators of Basic Service Access	Data Used	Brief Description	Key Finding on Inequality/Clustering
	21. Growing inequities in maternal health in South Africa...	Wabiri, N., et al.	2016	Longitudinal, Comparative Analysis	National household surveys (1998-2012)	Maternal health indicators (antenatal care, delivery).	Skilled birth attendance, antenatal care visits.	Serial South African Demographic and Health Survey (SADHS) microdata.	Tracks trends in maternal health service use and equity over time.	Found increasing socioeconomic inequality, with the poorest wealth quintiles forming persistent clusters of low service utilization.
	23. Clustering the E-Government services awareness... through K-Means algorithm	Malicaya, C.L., et al.	2021	Quantitative, Unsupervised ML (K-means)	LGU employee survey data	Awareness levels of e-gov services.	Awareness/access to information about e-government services.	Primary survey microdata of local government unit (LGU) employees.	Clusters local government units based on employee awareness of digital services.	Identified distinct clusters (High, Medium, Low awareness), highlighting inequality in internal readiness for digital service delivery.
	24. Household resilience and mitigating strategies to conflicts and shocks...	Muriuki, J., et al.	2024	Quantitative, Resilience Analysis	Household survey data (Uganda, Malawi)	Resilience capacity index, shocks, coping strategies.	Access to basic services (as a component of resilience), Access to Information (mobile phone).	Household survey microdata (e.g., World Bank LSMS-ISA).	Analyzes factors determining household resilience to shocks.	Households with low resilience capacity (linked to asset poverty) cluster together, indicating groups at systemic risk.

Theme	Article Title	Author(s)	Year	Research Design	Participants/Data Sources	Variables Observed	Indicators of Basic Service Access	Data Used	Brief Description	Key Finding on Inequality/Clustering
	25. Access to special support grants and relative impact of resilience pillars...	Olawuyi, S.O., & Mushunje, A.	2023	Quantitative, RIMA Analysis	Household survey data (Eastern Cape)	Resilience pillars (ABS, AST, SSN, AC).	Access to Basic Services (ABS) pillar: includes water, sanitation, health, education, electricity.	Household survey microdata analyzed using FAO's RIMA tool.	Uses FAO's RIMA tool to analyze household resilience capacity and its pillars.	Found unequal access to grants, and that the impact of resilience pillars varied, creating distinct household clusters based on their resilience structure.

5. REFERENCES

[1] Baba, Y., Arai, Y., Yasui, H., & Yonezawa, K. (2015). Area disparity of income and expenditure in Sri Lanka: Based on micro-data sets of the Household Income and Expenditure Survey of 2006–2007. In Poverty, Inequality and Growth in Developing Countries (pp. 128-152). Routledge.
[2] Balakrishnan, S., Srivastava, S., & Kothari, C. Empirical Evidence for Synergistic Influence of Regional Infrastructure Availability and Access on Community Vulnerability. Available at SSRN 5475572.
[3] Charrahy, Z., Aghamohammadloo, F., & Shamoradi Samani, R. (2025). Spatial Modeling and Analysis of Poverty with a Focus on Spatial Inequalities (Case Syudy: Zanjan Province). Town and Country Planning, 16(2), 285-301.
[4] David, A., Guilbert, N., Hamaguchi, N., Higashi, Y., Hino, H., Leibbrandt, M., & Shifa, M. (2018). Spatial poverty and inequality in South Africa: A municipality level analysis.
[5] Deitz, S., & Meehan, K. (2019). Plumbing poverty: mapping hot spots of racial and geographic inequality in US household water insecurity. Annals of the American Association of Geographers, 109(4), 1092-1109.
[6] Ekiru, M. N. (2020). Community Resilience Capacity, Environmental Factors and Performance of Food Security Projects in Loima Sub-county, turkana Kenya (Doctoral dissertation, University of Nairobi).
[7] Fadilah, Y. C., Sani, A., & Andrianingsih, A. (2025). APPLYING K-MEANS CLUSTERING FOR GROUPING PAPUA'S DISTRICTS BASED ON POVERTY INDICATORS ANALYSIS. JITK (Jurnal Ilmu Pengetahuan dan Teknologi Komputer), 10(3), 543-553.
[8] Gong, P. I. X. I. N., Huang, X. I. A. O. R. A. N., Huang, C. H. E. N. Y. U., & White, M. A. R. C. U. S. Quantifing the Imbalance of Spatial Distribution of Elderly Service with Muti-source Data. In Proceedings of the 27th Conference on Computer Aided Architectural Design Research in Asia (CAADRIA) (Vol. 1, pp. 455-464).
[9] Jiang, W., Yang, H., & Liu, C. (2025). Deeper Effects of fiscal multidimensional poverty reduction: household characteristics, financial lags and elite capture. PLoS One, 20(2), e0319255.
[10] Katumba, S. (2018). Spatial statistical analyses to assess the spatial extent and concentration of multidimensional poverty in Gauteng using the South African Multidimensional Poverty Index. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, 42, 85-92.
[11] Kibret, G. D., Demant, D., Dawson, A., & Hayen, A. (2024). Spatial patterns of maternal and neonatal continuum of care use and its correlations with women's empowerment. BMC Health Services Research, 24(1), 1018.
[12] Lopéz-Carr, A. C., Lopéz-Carr, D., Grant, L., & Weeks, J. (2017). Using spatial, hierarchical, and econometric models in urban data-poor areas to examine food security. Plurimondi, (17).
[13] Malicaya, C. L., Talirongan, H., Talirongan, F. J. B., & Sarend, J. G. (2021). Clustering the E-Government services awareness of the local government unit through K-Means algorithm. [Unpublished manuscript].
[14] McCordic, C., Frayne, B., Sunu, N., & Williamson, C. (2022). The household food security implications of disrupted access to basic services in five cities in the Global South. Land, 11(5), 654.
[15] McCordic, C., Frayne, B., & Sgro, J. (2020). No. 42: The Role of Infrastructure Access in Urban Household Vulnerability to Food Insecurity in Southern Cities.

[16] Muriuki, J., Hudson, D., & Fuad, S. (2024). Household resilience and mitigating strategies to conflicts and shocks: evidence from household data in Uganda and Malawi. Agrekon, 63(1-2), 65-81.
[17] Nguetse Tegoum, P., & Hevi, K. D. (2016). Child poverty and household poverty in Cameroon: a multidimensional approach. In Poverty and Well-Being in East Africa: A Multi-faceted Economic Approach (pp. 11-33). Cham: Springer International Publishing.
[18] Niranjan, V., Chouhan, A., Kushwaha, A., Gangwal, R., & Verma, M. (2025). Mapping of geographic inequality in mental health care facilities and psychiatrists distribution across seven districts of Indore division: A geospatial analysis from central India. Indian Psychiatry Journal, 34(4), 522–527.
[19] Noma, F., & Babu, S. (2024). Predicting climate smart agriculture (CSA) practices using machine learning: A prime exploratory survey. Climate Services, 34, 100484.
[20] Olawuyi, S. O., & Mushunje, A. (2023). Access to special support grants and relative impact of resilience pillars on households' resilience capacity in the Eastern Cape, South Africa: Application of FAO's Shiny RIMA tool.
[21] Oliveira, L. T. (2021). The Diversity of Deprived Areas: Applications of Unsupervised Machine Learning And Open Geodata.
[22] Parreno Jr, R. P., & Del Mundo, R. (2015). Clustering algorithm as A planning support tool for rural electrification optimization. Int. J. Sci. Technol. Res., 4(12), 112-120.
[23] Popoola, O., Adekoya, Y., & Obasanmi, T. (2016). Geospatial Analysis of the Distribution of Health Facilities in Peri-urban area of Ifedore, Nigeria.
[24] Saha, A., Hayen, A., Ali, M., Rosewell, A., MacIntyre, C. R., Clemens, J. D., & Qadri, F. (2018). Socioeconomic drivers of vaccine uptake: An analysis of the data of a geographically defined cluster randomized cholera vaccine trial in Bangladesh. Vaccine, 36(31), 4742-4749.
[25] Shifa, M., & Leibbrandt, M. (2024). Spatial inequality in sub-Saharan Africa. African Geographical Review, 43(1), 1-17.
[26] Sisay, D., Muche, T., & Ali, H. (2020). Spatial distribution and associated factors of institutional delivery among reproductive-age women (15-49 years) in Ethiopia: the case of Ethiopian Demographic and Health Survey 2016 Data.
[27] Wabiri, N., Chersich, M., Shisana, O., Blaauw, D., Rees, H., & Dwane, N. (2016). Growing inequities in maternal health in South Africa: a comparison of serial national household surveys. BMC pregnancy and childbirth, 16(1), 256.
[28] Waghmare, H., Maske, S., Thomas, N. T., Roy, S., & Gupt, J. (2025).Spatial Patterns and Developmental Determinants of Women's and Children's Nutrition in India.
[29] Wei, C., Cabrera-Barona, P., & Blaschke, T. (2016). Local geographic variation of public services inequality: does the neighborhood scale matter?. International journal of environmental research and public health, 13(10), 981.
[30] Zhang, Y., & Fan, Y. (2025). Equity evaluation of community-based public EV charging services–A case study of the Sacramento region. Energy Policy, 198, 114495.

Received: December 16, 2025
Accepted: February 6, 2026