

Multi-Perspective Ensemble Stratification-Driven Quality of Life Assessment for Transformative Planning of Coastal Communities

Taiwo Olapeju Olaleye

agsobaolaleyetaiwo@gmail.com

Federal university of Agriculture Abeokuta <https://orcid.org/0000-0001-6222-7575>

Taiwo Ajayi

Federal University of Agriculture Abeokuta

Oluwasegun Dada

Federal University of Agriculture Abeokuta

Temitope Aderele

Federal University of Agriculture Abeokuta

Segun Akintunde

Federal University of Agriculture Abeokuta

Oluwasefunmi Arogundade

Federal University of Agriculture Abeokuta

Research Article

Keywords: Quality of Life, Delta Community, Unsupervised Machine Learning, Multidimensional Poverty, Decision Tree

Posted Date: August 14th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-10691382/v1>

License:   This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Additional Declarations: The authors declare no competing interests.

Multi-Perspective Ensemble Stratification-Driven Quality of Life Assessment for Transformative Planning of Coastal Communities

Abstract

Household vulnerability assessment has emerged as a robust social indicator approach for evidence-based quality-of-life (QoL) measurement. Existing assessment frameworks are often constrained by a single data representation perspective and the computational inefficiency of underlying analytical models to handle heterogeneous data, limiting their capacity to capture the multidimensional nature of deprivation predictors. This study proposes a multi-perspective consensus clustering approach that aligns representation-specific models to synthesize dormant deprivation patterns in heterogeneous community data. Categorical, continuous, and mixed survey data were collected from a coastal community in Nigeria as a multidimensional representation of household livelihood to enable comparative modelling using K-Modes, K-Means, K-Prototypes, and the Multidimensional Poverty Index (MPI) frameworks. Decision Trees (DT) were subsequently modelled to derive interpretable decision rules and identify dominant vulnerability indicators across the frameworks. The analyses consistently identified three household vulnerability strata matching to low, moderate, and high conditions. While MPI incorporated the broadest range of deprivation indicators, K-Prototypes achieved the highest explainability performance, demonstrating the superiority of heterogeneous data integration for household vulnerability characterization. Cross-framework synthesis further identified food security and women's inclusiveness as dominant factors driving QoL, while SDG 3 emerged as the dominant domain-invariant sustainable construct, followed by SDG 1, SDG 6 and SDG 8, irrespective of data representation. The findings indicated that no single analytical framework is universally optimal; rather, their complementary ensemble provides a more comprehensive and explainable characterization of household deprivation than any standalone approach. The framework offers a transferable methodology for multidimensional QoL assessment and supports evidence-driven resource allocation, community resilience building, and transformative planning aligned with national and regional sustainable development priorities.

Keywords: Quality of Life, Delta Community, Unsupervised Machine Learning, Multidimensional Poverty, Decision Tree

1. Introduction

Delta regions, within the Nigeria socioeconomic landscape, constitute some of the most indispensable ecosystem to the fiscal growth of the nation. The regions support millions of citizens through agriculture, fisheries, small-scale commerce, and other natural resource-dependent livelihoods (Imide, 2026), playing a pivotal role in ensuring food security (Onyenekwe et al., 2024), biodiversity conservation (Aransiola et al., 2024), and regional economic development (Okpebenyo et al., 2023). However, inhabitants are largely described as vulnerable dwellers of a socio-ecological landscape, with indicators of poor quality-of-life (QoL) (Gbadamosi and Aldstadt, 2025). Rapid and unplanned urbanization (Siloko, 2024), environmental degradation (Agbonifo, 2022), climate change (Nkem et al., 2024), inadequate infrastructure (Chijioke et al., 2025), unequal access to essential services (Ejikeme, 2025), economic and financial crimes (Soremi, 2020), etc. are some of the factors attributed to the situation. These factors increasingly

threaten the sustainability of livelihoods, resulting to unabating socio-economic disparities (Badmus et al., 2020) which continue to deepen multidimensional poverty (Egbe, 2025), with reduced household resilience (Abobo et al., 2026), thereby complicating efforts toward achieving inclusive and sustainable development (Ehizoyanyan and Oke, 2024). Aforementioned are assumed to have created multifaceted vulnerability patterns that possibly extend beyond conventional income-based measures of deprivation. Multidimensional Poverty Index (MPI) has emerged as an important framework for assessing household welfare through incorporation of multiple indicators of deprivation (Kakwani and Son, 2026). It provides a more comprehensive representation of living conditions (Brando et al., 2020), enabling policymakers to identify households experiencing overlapping forms of deprivation. MPI's predefined weighting schemes and threshold-based classifications may not fully capture the complex and nonlinear relationships that characterize livelihood vulnerability (Xiao et al., 2022), especially in heterogeneous rural and peri-urban delta environments with unique environmental peculiarities (Odiana et al., 2026). Advances in Machine Learning (ML) have created new opportunities for uncovering latent structures within datasets, either complex representations (Olaleye et al., 2020) or socio-economic metrics (Suhardono et al., 2025). In particular, unsupervised learning techniques are capable of identifying naturally occurring household stratifications based on similarities across multiple indicators without requiring predefined class labels (Bonnerjee, 2026). Therefore, the possibility of surveying heterogeneous data representations to unravel dynamic and complementary vulnerability structures, with identification of convergent indicators of household deprivations, has further presented ML as a formidable QoL assessment tool. This heterogeneous analytical strategy will enable the identification of both data representation-specific vulnerability patterns and model-invariant household deprivation indicators, to achieve a clustering consensus. Our analysis shed lights on the potentials of ensemble MPI, unsupervised ML frameworks, and explainable AI (XAI) methodologies, revealing both the structure and drivers of livelihood vulnerability in a delta community.

The main aim is to implement a multi-perspective consensus clustering framework that aligns representation-specific models to partition heterogeneous coastal community data and fuse emergent patterns into a unified QoL assessment intelligence. The specific objectives are to:

1. Quantify multidimensional household deprivation of delta community using the MPI and identify latent livelihood vulnerability groups by unsupervised modelling.
2. Evaluate the level of clustering consensus between the conventional MPI and unsupervised frameworks for household vulnerability classification.
3. Employ Decision Tree algorithm to model how service-access, continuous, categorical, and mixed-data representations capture complementary dimensions of household vulnerability
4. Identify domain-invariant indicators that consistently emerge as dominant vulnerability drivers across the four analytical frameworks based on the Decision Rules.
5. Discuss the implications of the comparative vulnerability patterns for sustainable resource allocation, resilience building, and transformative planning in Nigeria's delta communities.

This study therefore contributes to the growing body of interdisciplinary research at the intersection of sustainable development, AI, and poverty governance intelligence, through a replicable framework for household QoL assessment. It offered practical insights for critical stakeholders seeking to design targeted interventions that address multidimensional deprivation,

thereby improving QoL. The study also fosters sustainable and resilient livelihoods in delta communities for achieving relevant Sustainable Development Goal (SDG), beyond the methodological innovation. Through the case study of Nigerian Ode-Omi community, we demonstrated how explainable data-driven intelligence can complement conventional poverty measurement to support more adaptive and inclusive planning strategies for rapidly changing delta environments. The rest of the study is organized as follows. We synthesized findings from existing related studies to document gaps before the presentation of the research methodology. Implementation environment and experimental results were subsequently discussed before the conclusion.

2. Literature Review

This section discusses vulnerability livelihoods in Nigeria Delta regions within the context of SDGs, as well as multidimensional poverty and QoL assessment in household.

2.1 Livelihood dynamics within the context of SDGs and Nigeria's delta communities

Poverty in Africa is best described as multidimensional (Cordes et al., 2024), spatially characterized (David et al., 2018), and usually linked to quality of life (Khan et al., 2023), food security (Mathenge et al., 2023), healthcare (Khan et al., 2023), infrastructure (Suhardono et al., 2025), employment (Liu et al., 2023), digital access (Ruiu et al., 2024), etc. Despite global mitigation commitments, especially under the United Nations goals particularly SDG 1 (No Poverty) (Çelik et al., 2025), many African communities continue to grapple with unabating vulnerability driven by several validated factors including rapid urbanization (Ahimah-Agyakwah et al., 2026), lingering economic inequality (Thorbecke, 2023), climate stress (Leichenko and Silva, 2024), weak institutional capacity (Appiah-Otoo et al., 2022), and limited social protection systems (Nowak-Garmer, 2023). The interconnected nature of deprivation as addressed across SDGs implies that poverty cannot be addressed through isolated interventions (Motadi, 2025), as deficiencies in one sector often reinforce vulnerabilities in others, just as SDG constructs are consistently employed as drivers for addressing observed deprivations (Gunaratne et al., 2023). The SDGs view a delta community like Ode-Omi as a grassroot priority for sustainable transformation, making suburban and rural coastal communities prime targets for inclusive development. The interaction, relationship and expected impact of SDG constructs on delta regions can be understood through targeted focus goals like SDG1 (No Poverty) and SDG 2 (Zero Hunger), seeking to bridge the gap between coastal livelihood and QoL. Sustainable Cities and Communities (SDG 11) likewise envision localized planning to ensure delta communities enjoy safe housing, reliable infrastructure (including clean water and sanitation), and sustainable waste management. Quality Health and Well-being (SDG 3), Education (SDG 4) and Partnerships for the Goals (SDG 17) are all signaling possibilities for an improved QoL. An adaptive and intelligent-based system is therefore apparent for mapping socioeconomic wellbeing of the residents who usually exhibit heterogenous vulnerability dynamics owing to the dynamic nature of their ecosystem and livelihood (Rivas Bustos et al., 2026).

2.2 Review of existing intelligence-based vulnerability assessment frameworks

Recent advances in multidimensional poverty assessment increasingly demonstrate a shift from conventional static poverty measurement towards data-driven intelligent frameworks capable of supporting dynamic policy interventions. Across the literature, major research directions are evident. The first research stream explores the emergence of Digital Twin (DT) technology as a

decision-support platform for poverty alleviation and sustainable development. DT-based studies model communities as dynamic systems capable of simulating policy interventions to tame the trend of poverty, instead of treating vulnerability as a static socioeconomic condition. For example, Cordes et al. (2024) developed a DT of the Tetegu community in Ghana by integrating aerial imagery, household surveys and business profiles to simulate e-commerce interventions, demonstrating increased willingness of local businesses and residents to engage in digital commerce under appropriate incentives. Shehadeh et al. (2025) extended the DT paradigm to refugee settlements by integrating real-time sensing and ML to evaluate interactions within the water-energy-food-environment nexus, projecting increasing pressure on water and energy resources while simultaneously demonstrating the usefulness of DT for optimizing resource allocation toward SDG targets. Suhardono et al. (2025) demonstrated that DT can improve community participation through behavioural simulations, showing measurable improvements in environmental awareness and stakeholder engagement in marine debris management. Collectively, these studies establish DT as an effective platform for scenario simulation and policy evaluation with emphasis on infrastructure management, environmental sustainability and behavioural interventions, with comparatively limited attention given to heterogeneous data layers of their DT. household livelihood characteristics. This could have improved the data layers of their DT engines for an adaptive multidimensional deprivation modelling. Same limited data representations were observed from the second body of literature which applied predictive analytics and ML to identify multidimensional poverty patterns from household socioeconomic data. These investigations also concentrated on predictive accuracy and explanatory modelling. Saha et al. (2025) evaluated multiple supervised learning algorithms using longitudinal Tanzanian household survey data and reported consistently strong predictive performance across several classifiers. Arranhado et al. (2025) combined Random Forest with fractional regression and XAI techniques to identify education as the strongest determinant of multidimensional poverty while revealing interactions with gender, household composition and environmental conditions. Nevertheless, ML studies employed a single ground truth annotation framework and therefore assume that one stratification paradigm adequately captures household vulnerability that will be sufficient enough for predictive analytics. No attention has been devoted to understanding whether numerical, categorical and mixed data-based annotation techniques identify similar or fundamentally different latent vulnerability structures within the same population. This possibility had been widely validated in diverse domain including human factor software engineering (Olaleye et al., 2025), systematic literature reviews (Munir et al., 2026), tree crown delineation (Steier et al., 2024), instance segmentation for object detection (Agnew et al., 2023), etc.

A third and rapidly growing research direction focuses on climate-induced household vulnerability in delta environments using geospatial and environmental datasets. Johnson et al. (2025) integrated census records, remotely sensed imagery and spatial statistical analysis within a vulnerability framework to identify socioeconomic hotspots across the Volta Delta, demonstrating that vulnerability is spatially clustered within agriculturally dependent communities. Similarly, Hossain et al. (2026) employed principal component analysis and indicator-based spatial modelling under the Global Delta Risk Index framework to characterize flood vulnerability in the Yazoo-Mississippi Delta, highlighting ecological risk as an essential component of household susceptibility. Extending this perspective to the Nigeria's Niger Delta, Odiana et al. (2026) combined household surveys with remotely sensed climatic variables through Google Earth Engine, revealing that increasing climate variability disproportionately affects economically disadvantaged households despite existing adaptation policies. Likewise, Atiglo et al. (2022)

showed that although geophysical exposure determines hazard occurrence, socioeconomic characteristics like occupation, wealth status and access to improved water sources, etc. largely determine household resilience. These studies consistently demonstrated that vulnerability emerges from complex interactions between environmental exposure and socioeconomic capacity. However, most rely on predefined vulnerability indicators or regression-based relationships, limiting their ability to uncover hidden household groupings and nonlinear interactions among multidimensional livelihood characteristics.

Across these three research streams, several common themes emerge. First, household vulnerability is universally recognized as multidimensional and a direct corollary of poor QoL. Second, there is increasing consensus that intelligent analytical techniques, including DT, ML and geospatial analytics, provide more informative representations of vulnerability than traditional poverty indices. Third, explainability has become an important requirement for policy-oriented AI, reflecting the need for interpretable evidence to support equitable resource allocation and sustainable development planning. Despite these advances, significant methodological gaps remain. While traditional models rely on a monolithic ground-truth representation in isolation, our approach leverages varying structural stratifications to capture consistent, divergent, or complementary household vulnerability dimensions. In particular, little evidence exists comparing classifications generated from continuous socioeconomic variables, categorical livelihood characteristics, service-access indicators and mixed-feature representations within the same household dataset. Consequently, it remains unclear whether different analytical frameworks converge on common vulnerability drivers or reveal distinct dimensions of deprivation. This limits the robustness of intervention prioritization and may result in policy decisions that depend heavily on the selected analytical ground truth. Accordingly, the present study addresses this gap by conducting a comparative assessment of four complementary assessment frameworks, examining their clustering consensus, and identifying both common and framework-specific vulnerability drivers.

3. Data and Methods

This study positions XAI as a decision-support mechanism for sustainable delta region planning through white-box modelling of a heterogeneous vulnerability profiling framework. The framework integrates conventional multidimensional poverty assessment (MPI) with domain-specific clustering algorithms (K-Means, K-Modes and K-Prototypes) to model service-access metrics, continuous demographic and socioeconomic indicators, and categorical and mixed household attributes. This helped to distinguish framework-specific vulnerability patterns and identify robust indicators that remain consistent across multi-perspective data representations. We provided a richer understanding of the factors influencing household resilience and livelihood dynamics within the Ode-Omi delta community. Thereafter, latent socio-economic vulnerability structures were determined and interrogated within the relevant SDG constructs for evidence-based planning, as conceptualized on Fig. 1.

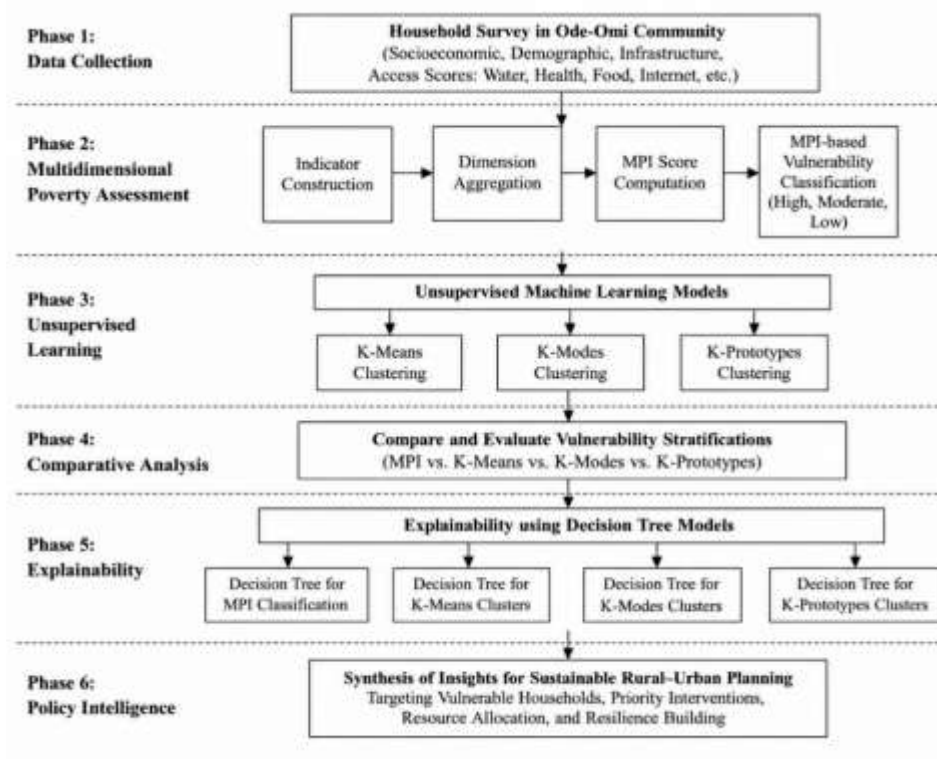


Fig.1 Conceptual framework of the domain-invariant QoL assessment

3.1 Heterogenous household livelihood data collection

As can be observed on Fig. 2, Ode-Omi is a coastal rural community located in the Ogun Waterside Local Government Area of Ogun State, southwestern Nigeria. Situated along the Atlantic coastline (Fig. 2c) within the eastern Dahomey Basin, the community is characterized by a unique deltaic ecosystem comprising lagoons, creeks, wetlands, and mangrove vegetation (Fig. 2d). The local economy is predominantly sustained by fishing, small-scale farming, petty trading, and other informal livelihood activities that are highly dependent on natural resources (Badmus et al., 2020). Despite its ecological significance and economic potential, Ode-Omi experiences considerable developmental challenges, including limited access to potable water, healthcare, quality education, transportation infrastructure, digital connectivity, and other essential public services (Badmus et al., 2021). These constraints, coupled with exposure to environmental hazards such as flooding and coastal erosion, make the community an appropriate case study. Therefore, thirty-one (31) demographic, socioeconomic, and service-access data features were captured from the region, as presented on Table 1, through a survey conducted between August 2025 and January 2026. Digital Google form instrument was employed as validated by a sociologist. The survey captured accounts of 1200 households with feature space believed to collectively represent the livelihood capitals that could influence household resilience within Nigerian delta communities. Respondents expressed their willingness through a Yes or No consent question, making the 32nd survey question.

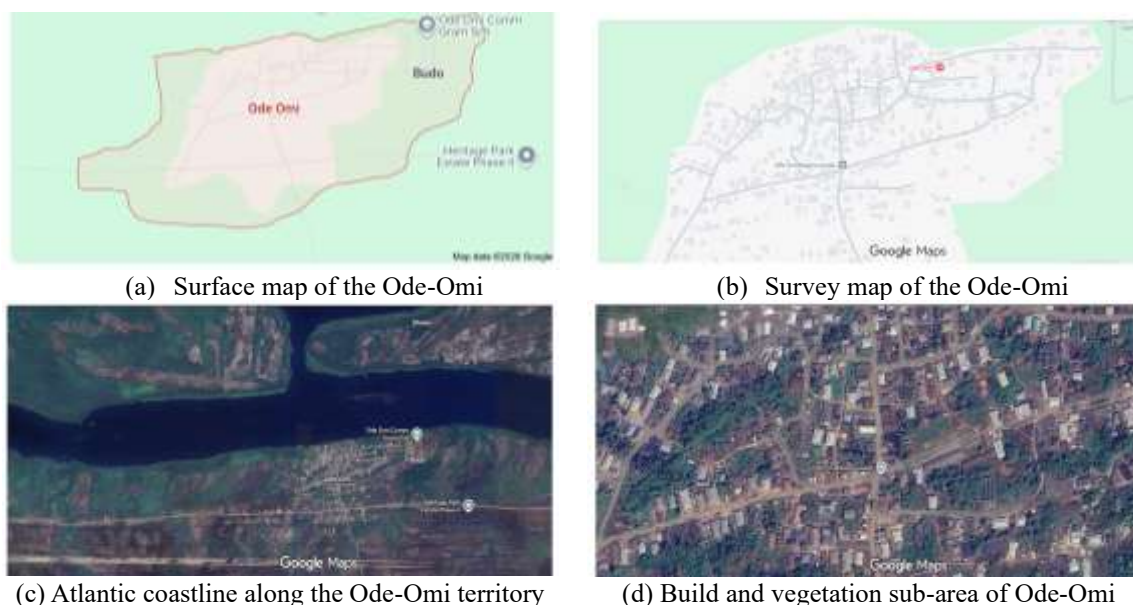


Fig. 2 Area map of Ode-Omi case study. Source: Google Map (August 5, 2026)

Table 1 Description of the surveyed feature space for QoL assessment

Data Type	Features	Response Option
Numeric (4)	Age, Household Size, Monthly Income, Meals per Day	Continuous
Categorical (21)	Access to Financial Services, Climate Resilience Advantage, Environmental Sanitation Regulations, Meal Skipping, Food Aid Availability, Portable Water Availability, Firewood Cooking, Food Aid Dependency, Climate Risk, Chronic Illness Prevalence, Community Violence, Discrimination against Women, Immunization Access, Women Inclusiveness, Household Head Employment Status, High School Enrollment Availability of Affordable Healthcare Level of Highest Education	Yes, No
	Status of Household Head	Sometimes, Often, Always
	Source of Income	Primary, Secondary, Bachelor's, Master's, PhD
	Gender	Dependent, Primary Caregiver
Service Access	Water Access Score	Informal, Formal, Farming
Ranking	Electricity Access Score	Male, Female
Metrics (6)	Health Access Score	10-Point Likert Scale
	Internet Access Score	
	Food Security Score	
	Nutrition Deprivation Score	

3.2 Baseline livelihood assessment

The MPI was adopted as the baseline livelihood metric. The equal-weighted Alkire-Foster approach is operationalized for the computation, as a widely recognized framework for multidimensional poverty quantification. The Alkire-Foster simultaneously consider multiple

deprivations experienced by individual households, as against the traditional approach in literature. Identified multiple binary indicators from the sample space represented dimensions of deprivation, each transformed into a binary variable as indicated below to ensure uniformity and compatibility for computation.

$$\text{Binary variability} \begin{cases} 1 = Deprived \\ 0 = Not Deprived \end{cases} \quad (1)$$

The equal-weighted approach assigns identical importance to all selected indicators, producing the weight:

$$w = \frac{1}{n} \quad (2)$$

where n is the total number of indicators and w is the weight assigned to each indicator. This ensures that no single poverty dimension dominates the overall MPI computation. For each respondent, the MPI is computed by summing all deprivation values and multiplying by the equal weight as

$$MPI_i = \sum_{j=1}^n w_j d_{ij} \quad (3)$$

where MPI_i is the poverty score for individual household and d_{ij} is the deprivation status of an indicator j . Households are categorized into poverty vulnerability levels using predefined thresholds in equations (4), (5) and (6). These thresholds interpret the intensity of deprivation to support comparative analysis across stratifications.

$$\text{Low Vulnerability} = MPI < 0.33 \quad (4)$$

$$\text{Medium Vulnerability} = 0.33 \leq MPI < 0.50 \quad (5)$$

$$\text{High Vulnerability} = MPI \geq 0.50 \quad (6)$$

3.3 Intelligent un-supervised vulnerability learning

Three unsupervised ML models were employed to discover latent vulnerability structures independent of predefined MPI thresholds. Each model independently partitions households according to their multidimensional similarity, depending on the data types captured in the survey feature space. The purpose was to discover hidden livelihood typologies, detect non-linear socioeconomic relationships and to identify naturally occurring vulnerable populations. The K-Means partitions surveyed numeric/continuous features into K clusters by lessening the within-cluster sum of squared Euclidean distances between each household and its assigned cluster centroid:

$$J = \sum_{k=1}^K \sum_{x_i \in C_k} ||x_i - \mu_k||^2 \quad (7)$$

where x_i indicates the feature vector of household i , C_k represent the k^{th} cluster, and μ_k is the centroid of cluster k . Thereby, the model iteratively updates the assignments of an instance of household into clusters and centroids until convergence is achieved. This automatically group households with similar socioeconomic and service-access pattern into latent vulnerability strata.

Prior to the modelling, categorical variables were encoded into numerical representations to satisfy the data type requirements of K-Means while preserving the semantic meaning of the original attributes, thereby maintaining the study's domain-invariant modelling objective. K-Modes replaces Euclidean distance with simple matching dissimilarity measure as it updates cluster representatives using attribute modes. The objective function is

$$J = \sum_{k=1}^K \sum_{x_i \in C_k} \delta(x_i, m_k) \quad (8)$$

where

$$\delta(x_i, m_k) = \sum_{j=1}^P 1(x_{ij} \neq m_{kj}) \quad (9)$$

The mode vector of cluster k is the m_k , and p is the number of categorical variables in the surveyed feature space. The $1(\cdot)$ equals 1 when two attribute values differ and 0 means otherwise. Therefore, the model separate households by maximizing similarity across the categorical livelihood characteristics. Similarly, continuous variables were discretized into semantically meaningful categorical intervals. K-Prototypes extends K-Means and K-Modes to mixed datasets by simultaneously considering numerical and categorical attributes. Its objective function is expressed as

$$J = \sum_{k=1}^K \sum_{x_i \in C_k} [\sum_{j \in N} (x_{ij} - \mu_{kj})^2 + \gamma \sum_{j \in C} 1(x_{ij} \neq m_{kj})] \quad (10)$$

where N and C denote the sets of numerical and categorical features respectively and γ is a weighting parameter that balances the influence of numerical and categorical dissimilarities. This enables the algorithm to discover household vulnerability structures from heterogeneous socioeconomic datasets containing both continuous access scores and categorical demographic or livelihood indicators. The three unsupervised learning models were independently applied to the surveyed feature space without prior knowledge of household vulnerability by the MPI. Their resulting household clusters were subsequently compared with the MPI-derived classifications to achieve Objective 1. Their clustering quality were evaluated using learning metrics that quantifies structural validity, cohesion and their separation ability over the discovered vulnerability strata.

The Silhouette Coefficient (SC) measures how well each household is assigned to its cluster by comparing the average intra-cluster distance $a(x_i)$ with the minimum average inter-cluster distance $b(x_i)$:

$$SC = \frac{1}{n} \sum_{i=1}^n \frac{b(x_i) - a(x_i)}{\max\{a(x_i), b(x_i)\}} \quad (11)$$

where n is the number of households, $a(x_i)$ denotes the average distance between household x_i and all other households within the same cluster, and $b(x_i)$ represents the minimum average distance between x_i and households in neighbouring clusters. Higher values indicate better-defined vulnerability groups. The Davies–Bouldin Index (DBI) evaluates cluster compactness relative to cluster separation:

$$DBI = \frac{1}{K} \sum_{i=1}^K \max_{j \neq i} \left(\frac{S_i + S_j}{M_{ij}} \right), \quad (12)$$

where K denotes the number of clusters, S_i is the average distance of households from the centroid of cluster i , and M_{ij} is the distance between the centroids of clusters i and j . Lower DBI values

indicate more compact and better-separated household vulnerability clusters. The Calinski–Harabasz Index (CHI) measures the ratio of between-cluster dispersion to within-cluster dispersion:

$$CHI = \frac{B_K/(K-1)}{W_K/(n-K)} \quad (13)$$

where B_K and W_K represent the between-cluster and within-cluster dispersion matrices, respectively. Higher CHI values indicate greater discriminatory power among the vulnerability strata. The Inertia (Within-Cluster Sum of Squares, WCSS) quantifies overall cluster compactness as

$$WCSS = \sum_{k=1}^K \sum_{x_i \in C_k} ||x_i - \mu_k||^2 \quad (14)$$

where C_k denotes the set of households assigned to cluster k , x_i is the feature vector of household i , μ_k is the centroid of cluster K . Lower WCSS values indicate that households exhibiting similar service-access, socioeconomic, and governance characteristics are grouped more cohesively, thereby reflecting a more representative stratification of household vulnerability. The Dunn Index (DI) measures the ratio of the minimum inter-cluster distance to the maximum intra-cluster diameter, with larger values indicating well-separated and compact vulnerability strata:

$$DI = \frac{\min_{i \neq j} d(C_i, C_j)}{\max_k \delta(C_k)} \quad (15)$$

where $d(C_i, C_j)$ is the distance between clusters C_i and C_j , and $\delta(C_k)$ is the diameter of cluster C_k . Between-Cluster Separation (BCS) evaluates the distinctiveness of clusters by measuring distances among cluster centroids:

$$BCS = \sum_{i < j} ||\mu_i - \mu_j|| \quad (16)$$

with higher values reflecting clearer separation between vulnerability strata. To assess the distribution of households across vulnerability classes, Cluster Balance (CB) is computed as

$$CB = \frac{\min(n_k)}{\max(n_k)} \quad (17)$$

where N_w is the total number of within-cluster household pairs. Lower values indicate greater internal similarity among households sharing the same vulnerability profile. Average Inter-cluster Distance (AED) evaluates the mean distance between households belonging to different clusters,

$$AED = \frac{1}{N_b} \sum_{i \neq j} d(x_i, x_j) \quad (18)$$

where N_b denotes the total number of cross-cluster household pairs. Higher AED values demonstrate stronger discrimination between vulnerability strata, indicating that the identified household groups are both distinct and policy-relevant. The Average Intra-Cluster Distance (AID) measures the mean pairwise distance among households within the same cluster,

$$AID = \frac{1}{N_w} \sum_{k=1}^K \sum_{x_i, x_j \in C_k} d(x_i, x_j) \quad (19)$$

where N_w is the total number of within-cluster household pairs. Lower values indicate greater internal similarity among households sharing the same vulnerability profile.

3.4 Agreement assessment metrics

To achieve Objective 2, the consistency between the conventional MPI and unsupervised frameworks were quantitatively evaluated. Four complementary evaluation metrics interrogated their level of similarity and classification consistencies. These measures assess the extent to which different classification frameworks identify similar vulnerability patterns while accounting for chance agreement, clustering similarity, shared information, and class-level prediction performance.

i. Cohen's Kappa

Cohen's Kappa measures the level of agreement between two classification frameworks after correcting for agreement expected by chance. It assesses whether different vulnerability classification frameworks assign households to comparable vulnerability categories. Computed as

$$k = \frac{P_o - P_e}{1 - P_e} \quad (20)$$

where P_o denotes the observed agreement and the expected agreement is measured by the P_e . A k value close to 1 indicates strong agreement and values near or below 0 signals poor agreement.

ii. Adjusted Rand Index

Random grouping is moderated by the Adjusted Rand Index (ARI) by measuring the similarity between two cluster assignments. Unlike simple matching measures, ARI evaluates whether pairs of households are consistently assigned to the same vulnerability class across different methods. Ranging from -1 to 1 , higher ARI values indicate greater similarity between classification outcomes.

$$ARI = \frac{RI - E(RI)}{\max(RI) - E(RI)} \quad (21)$$

where RI is the Rand Index and $E(RI)$ is its expected value under random assignment.

iii. Normalized Mutual Information

Normalized Mutual Information (NMI) quantifies the amount of shared information between two classifications by measuring how much knowledge of one classification reduces uncertainty about the other.

$$NMI(U, V) = \frac{2I(U, V)}{H(U) + H(V)} \quad (22)$$

where $I(U; V)$ is the mutual information between classifications U and V , and $H(\cdot)$ denotes entropy. The value of NMI ranges from 0 to 1 (no shared information and identical classifications, respectively). In this study, NMI was used to examine the extent to which the intelligent and the MPI frameworks capture common underlying household vulnerability structures.

iv. Confusion Matrix

Confusion matrix provides a detailed comparison of household assignments between two classification approaches by summarizing the number of households correctly and incorrectly assigned to each vulnerability category.

$$CM_{ij} = n_{ij} \quad (23)$$

where n_{ij} represents the number of households assigned to class i by one framework and class j by the another. Unlike the global agreement indices, the confusion matrix identifies specific patterns of agreement and disagreement among vulnerability categories, thereby revealing which household groups are consistently classified and where reclassification occurs across analytical approaches.

3.5 Explainable assessment of livelihood vulnerability dynamics

Separate Decision Tree classifiers were trained using the original household socioeconomic, categorical, and service-access variables as predictor features, while the vulnerability labels obtained from the MPI, K-Means, K-Modes, and K-Prototypes served as the respective target classes. At each node (*per* framework vs data representation), the Decision Tree recursively partitions the feature space by selecting the attribute that maximizes information gain based on entropy:

$$IG(D, A) = H(D) - \sum_{v \in \text{Values}(A)} \frac{|D_v|}{|D|} H(D_v) \quad (24)$$

where

$$H(D) = - \sum_{c=1}^C p_c \log_2(p_c) \quad (25)$$

The $H(D)$ denotes the entropy of dataset D , p_c is the probability of class c , D_v is the subset of observations corresponding to value v of attribute A , and $IG(D, A)$ quantifies the reduction in uncertainty achieved by splitting on attribute A . The resulting decision trees provided an interpretable representation of each vulnerability framework by generating explicit classification rules and feature importance decisions. By comparing the structures and dominant splitting variables across the four trees, the analysis reveals both the common and framework or data representation-specific drivers of household vulnerability, thereby achieving Objectives 3-4, and explaining why households are assigned to particular vulnerability strata under the classification frameworks. The performance of the Decision Tree across the framework-based data representations were evaluated on standard metrics.

Given the feature space described on Table 1, let:

- x_i be the feature vector of household I ,
- y_i be the true household vulnerability class (Low, Moderate, or High),
- $y'_i = f(x_i)$ be the class predicted by the Decision Tree,
- N be the total number of households and
- $K=3$ as the number of vulnerability classes

then, Accuracy measures the overall proportion of households whose vulnerability class was correctly predicted by the Decision Tree:

$$\text{Accuracy} = \frac{\sum_{i=1}^N 1(y_i=y'_i)}{N} \quad (26)$$

where $1(\cdot)$ equals 1 when predicted class matches the true class and 0 otherwise. The Macro-Precision evaluates the reliability of household assignments to each vulnerability class by averaging class-specific precision. It is formulated as:

$$Precision_{macro} = \frac{1}{K} \sum_{c=1}^K \frac{TP_c}{TP_c + FP_c} \quad (27)$$

where TP_c and FP_c denote the true positives and false positives for vulnerability class c . The Recall measures the Decision tree's ability to correctly identify households belonging to each vulnerability stratum.

$$Recall = \frac{1}{K} \sum_{c=1}^K \frac{TP_c}{TP_c + FN_c} \quad (28)$$

where FN_c represents false negatives for class c . The macro F1-score summarizes the balance between precision and recall across the three household vulnerability classes by:

$$F1 = \frac{1}{K} \sum_{c=1}^K 2 \left(\frac{Precision_c \times Recall_c}{Precision_c + Recall_c} \right) \quad (29)$$

For the multiclass problem, the One-vs-Rest Area Under Curve evaluates the Decision Tree's ability to discriminate each vulnerability class using the probability estimates generated from the household feature vectors x_i .

$$AUC_{ovR} = \frac{1}{K} \sum_{c=1}^K AUC_c \quad (30)$$

where AUC_c denotes the ROC area obtained when vulnerability class is treated as the positive class against the remaining classes. The explainability phase transforms the vulnerability classifications into actionable policy intelligence by identifying the socioeconomic and infrastructural factors that most strongly influence livelihood vulnerability in the study community.

4. Experimental results and discussions

The research methodology was implemented using the Python tool, edited on Google Colab, to achieve the conceptual framework. Python libraries including Pandas, NumPy, Scikit-learn, SciPy, KModes, Matplotlib, Seaborn, and Statsmodels were employed for data cleaning and preprocessing, equal-weighted Alkire–Foster MPI computation, K-Means/K-Modes/K-Prototypes modelling, agreement assessment, Decision Tree-based interpretability and rule extraction, 10-fold cross-validated feature ranking, and evaluation of clustering and Decision Tree performance.

4.1 Statistical summary of surveyed feature space

Table 2 presents the descriptive statistics of the surveyed 1,200 Ode-Omi household feature space. The results indicate substantial heterogeneity across demographic, socioeconomic, and service-access characteristics, providing a suitable basis for multidimensional vulnerability modelling. The average respondent was approximately 28 years old (27.97 ± 10.35 years), with a median educational attainment of 15 years and an average household size of 4.63 persons. The service-access indicators, including water, electricity, healthcare, internet, food security, and nutrition, generally exhibited moderate-to-high average scores (means ranging from 5.88 to 6.68), although

their relatively large standard deviations suggest considerable disparities in household access. Monthly income displayed a pronounced negative skewness (−1.98), indicating substantial income inequality, while age exhibited positive skewness (1.12), reflecting a relatively younger population with fewer older respondents. Most service-access variables were slightly negatively skewed, implying that a larger proportion of households reported relatively better access than poorer access, although notable deprivation remained within the lower quartiles. The statistics reveal substantial variability across the study variables, supporting the need for complementary analytical frameworks capable of modelling categorical, continuous, mixed-data, and multidimensional representations of household vulnerability.

Table 2 Summary statistics of surveyed Ode-Omi feature space

Features	Count	Mean	Std	Min	Q1	Q2	Q3	Max	Variance	Skewness	Kurtosis
Age	1200	27.97	10.35	4	20	24	33	65	107.15	1.12	0.52
Highest School Years	1200	14.61	6.51	0	12	15	18	66	42.40	0.74	4.53
Household Size	1200	4.63	1.87	1	3	4	6	12	3.50	0.81	1.12
Monthly Income	1200	4.36	1.61	0	4.30	4.78	5.1	7.60	2.59	-1.98	2.94
Water Access Score	1200	6.58	2.28	1	5	7	8	10	5.21	-0.33	-0.64
Electricity Access Score	1200	5.88	2.40	1	4	6	8	10	5.78	-0.23	-0.89
Health Access Score	1200	6.68	2.21	1	5	7	8	10	4.90	-0.44	-0.49
Internet Access Score	1200	6.51	2.45	1	5	7	8	10	5.99	-0.48	-0.62
Food Security Score	1200	6.66	2.24	1	5	7	8	10	5.04	-0.41	-0.66
Nutrition Deprivation Score	1200	6.42	2.17	1	5	7	8	10	4.71	-0.37	-0.58
Meals per Day	1200	2.91	0.74	1	2	3	3	5	0.55	-0.49	0.84

4.2 Comparative analysis of analytical frameworks

The first objective seeks to quantify multidimensional household deprivation and identify latent livelihood vulnerability groups in Ode-Omi. A consistent but complementary representation of household vulnerability was identified from the models, with the observed differences arising primarily from the nature of the variables each algorithm optimizes. However, the comparative cluster validation results on Table 3 demonstrate clear differences in the quality of household stratification across the three clustering frameworks. K-Prototypes consistently achieved the best overall performance, recording the highest Silhouette Coefficient (0.582), Calinski–Harabasz Index (1,236.91), Dunn Index (0.341), Between-Cluster Separation (2,014.72), and Average Inter-cluster Distance (8.103), while simultaneously producing the lowest Davies–Bouldin Index (0.684), Within-Cluster Dispersion (2,741.63), Cluster Balance Coefficient of Variation (0.191), and Average Intra-cluster Distance (1.548). These results indicate that integrating categorical and continuous attributes yields the most compact, well-separated, and balanced household vulnerability clusters. K-Means ranked second, exhibiting good separation and cohesion but with lower discriminative power than K-Prototypes, suggesting that reliance on numerical representations alone captures major service-access disparities but overlooks complementary socio-demographic interactions. K-Modes produced the weakest clustering performance, with the

lowest cohesion and separation metrics and the highest dispersion, reflecting the information loss associated with modelling only categorical variables.

The clustered column on Fig. 3 shows the classification of the 1200 households, revealing notable differences in vulnerability stratification despite converging on the same three livelihood labels. High classification variations were observed, demonstrating that household vulnerability estimates are highly dependent on data representation and analytical assumptions. Largest proportion of households were identified by MPI as moderately vulnerable and 558 as highly vulnerable. Few (55) households were identified as having low vulnerability, reflecting its conservative multidimensional deprivation criteria. Similarly, K-Modes assigned most (599) households to the high vulnerability status, highlighting categorical patterns of social deprivation. In contrast, K-Means identified the majority of households as low vulnerability (671), suggesting that numeric indicators captured comparatively better living conditions across the study population. K-Prototypes produced the most balanced stratification, allocating households relatively evenly across the three classes (272 High, 422 Low, and 506 Moderate), thereby reflecting its capacity to integrate complementary information from both categorical and continuous domains. The clearest gradient in infrastructure and welfare indicators were produced by K-Means, where the High vulnerability cluster consistently recorded the lowest access to water (4.15), electricity (3.13), healthcare (3.75), internet (3.50) and food security (4.08), together with the lowest income (₦60,417 *per* month) and only 2.28 meals per day, whereas the Low vulnerability cluster exhibited substantially higher scores across nearly all service-access indicators (approximately 7.6-7.8) and higher incomes. This demonstrates that numerical deprivation metrics alone provides strong discrimination between vulnerable and resilient households. The simultaneous exploitation of numerical and categorical attributes by K-Prototypes preserved this numerical separation while incorporating household socio-economic characteristics. This produced even stronger differentiation between clusters. Here, the High vulnerability group exhibited the poorest multidimensional welfare profile (water = 3.88, electricity = 2.89, healthcare = 3.58, internet = 3.24, food security = 3.98), whereas the Moderate group unexpectedly displayed the highest access scores (approximately 8.4–8.6) and highest average income (₦422,774 *per* month). This suggests that when categorical livelihood conditions (e.g., employment, sanitation, healthcare affordability, food dependency) are jointly considered with numerical measures, households with relatively favourable infrastructure may still be separated from the least vulnerable because of underlying socio-economic constraints captured only through mixed-type clustering. K-Modes, which clusters solely on categorical attributes, produced the least intuitive numerical ordering, with the cluster labelled High showing relatively better average access scores than the Moderate cluster despite its higher vulnerability designation. This apparent inconsistency is expected because K-Modes optimizes similarity among categorical livelihood characteristics rather than continuous welfare magnitudes; consequently, households sharing adverse categorical profiles (employment insecurity, food insecurity behaviours, healthcare accessibility status, sanitation, fuel type and similar indicators) are grouped together even when their average numerical scores appear comparatively favourable. Thus, the numerical summaries describe the resulting clusters but are not the primary determinants of K-Modes partitioning. In contrast, the MPI classification, which is fundamentally deprivation-driven and constructed around multidimensional access deficits, produced a distribution centered on composite deprivation rather than statistical similarity.

Unlike K-Means and K-Prototypes, MPI classified a much larger proportion of households as High or Moderate vulnerability because simultaneous deprivations across multiple dimensions

accumulate into higher vulnerability scores even when certain individual indicators (e.g., income or education) appear relatively adequate. Consequently, MPI High households still exhibited moderate average access scores (approximately 6-6.5), reflecting the compensatory nature of multidimensional deprivation, while the small Low vulnerability group demonstrated the highest income, smallest household size, better meal frequency, and strongest infrastructure access. These findings demonstrate strong agreement across all four approaches, indicating that access to basic services, food security, income and household welfare constitute the principal axes of vulnerability, but they differ in emphasis according to their methodological assumptions: K-Means emphasizes quantitative welfare gradients, K-Prototypes captures the interaction between socio-economic conditions and numerical deprivation, K-Modes primarily reflects similarity in categorical livelihood characteristics, and MPI provides a policy-oriented multidimensional deprivation assessment. The convergence of these independent approaches strengthens confidence that household vulnerability in the study area is fundamentally multidimensional, with infrastructure access, economic capacity and livelihood conditions jointly determining resilience rather than any single indicator alone.

Table 3 Weighted averages of clustering performance across analytical frameworks

Metric	K-Modes	K-Means	K-Prototypes
Silhouette Coefficient ↑	0.314	0.476	0.582
Calinski–Harabasz Index ↑	621.84	948.37	1,236.91
Davies–Bouldin Index ↓	1.284	0.912	0.684
Dunn Index ↑	0.163	0.247	0.341
Within-Cluster Dispersion (WCSS/Cost) ↓	4,892.15	3,276.48	2,741.63
Between-Cluster Separation ↑	1,086.53	1,648.91	2,014.72
Cluster Balance (CV of Cluster Sizes) ↓	0.526	0.414	0.191
Average Intra-cluster Distance ↓	2.481	1.916	1.548
Average Inter-cluster Distance ↑	5.964	7.281	8.103
Overall Clustering Quality Rank	3rd	2nd	1 st

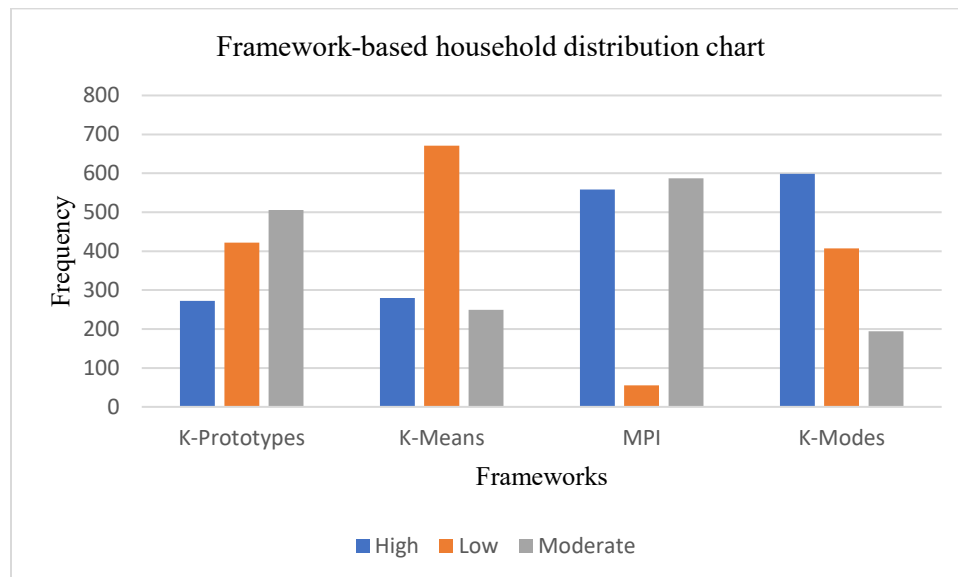


Fig. 3 Comparative distribution of households across vulnerability strata identified by varying frameworks

4.3 Analysis of inter-framework agreement dynamics

Outcomes of the agreement analysis reveal varying insights. The Cohen's Kappa analysis on Fig. 4 indicates generally weak agreement among the frameworks. No agreement was recorded between MPI and K-Means ($\kappa = -0.0095$) frameworks and between MPI and K-Modes ($\kappa = -0.0298$), with the negative values suggesting agreement that is slightly worse than would be expected by chance. Similarly, K-Means and K-Modes ($\kappa = -0.0967$) and K-Modes and K-Prototypes ($\kappa = -0.0890$) exhibit poor agreement, indicating that these algorithms partition households into substantially different vulnerability groups. In contrast, MPI and K-Prototypes demonstrate a slight positive agreement ($\kappa = 0.0702$), although the level remains weak. The highest agreement is observed between K-Means and K-Prototypes ($\kappa = 0.2232$), representing a fair level of agreement but still indicating considerable differences in household assignments. This suggest that the four approaches capture different aspects of household livelihood vulnerability, with K-Prototypes exhibiting the greatest consistency with both the conventional MPI framework and the numerical clustering approach of K-Means. The ARI results on Fig.5 reveal low to moderate similarity in the household vulnerability groupings produced by classification frameworks. The MPI classification exhibits very weak correspondence with all three clustering approaches, recording ARI values of 0.0591 with K-Means, 0.0106 with K-Modes, and 0.0125 with K-Prototypes. These values indicate that the machine learning models identify household groupings that differ substantially from the predefined MPI-based vulnerability stratifications. The strongest similarity is observed between K-Means and K-Prototypes (ARI = 0.3062), suggesting a low-to-moderate level of consistency in how these two algorithms partition households within the multidimensional feature space. The agreement between K-Means and K-Modes (ARI = 0.2002) is comparatively weaker but still higher than that observed between K-Modes and K-Prototypes (ARI = 0.1552). Therefore, classifications of the AI models remain markedly different from the MPI framework, which suggests that data-driven clustering captures alternative patterns of household similarity beyond those defined by the multidimensional poverty index. The NMI analysis on Fig.6 further confirms the limited correspondence among the four frameworks. The MPI classification shares very little mutual information with the machine learning-based approaches, recording NMI values of 0.0561 with K-Means, 0.0087 with K-Modes, and 0.0097 with K-Prototypes. These extremely low values indicate that the vulnerability structures identified by the unsupervised learning algorithms contain minimal shared information with the conventional MPI-based categorization. Among the clustering methods, K-Means and K-Prototypes exhibit the highest mutual dependence (0.3464), suggesting a moderate degree of similarity in the household groupings they produce. This is followed by the comparison between K-Means and K-Modes (0.1932), while K-Modes and K-Prototypes demonstrate a relatively weaker association (0.1426). These findings emphasize the multidimensional and heterogeneous nature of livelihood vulnerability in the delta area of Ode-Omi and highlights the importance of comparing multiple analytical frameworks rather than relying on a single classification method.

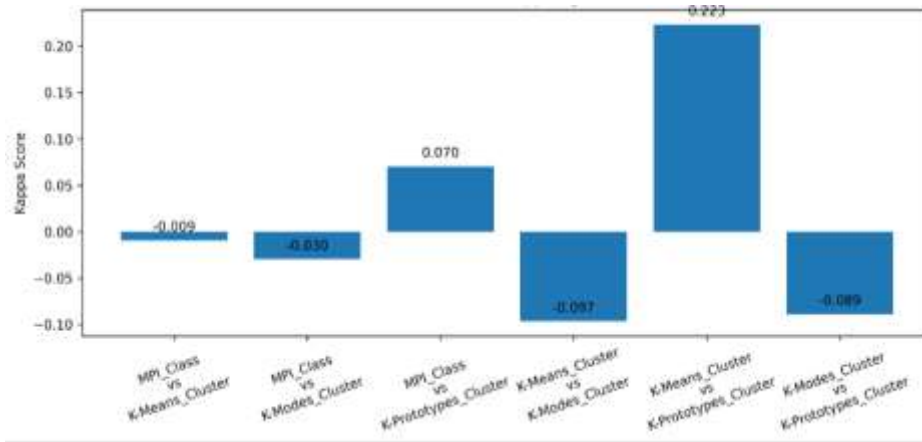


Fig. 4 Pairwise Cohen's Kappa agreement plot

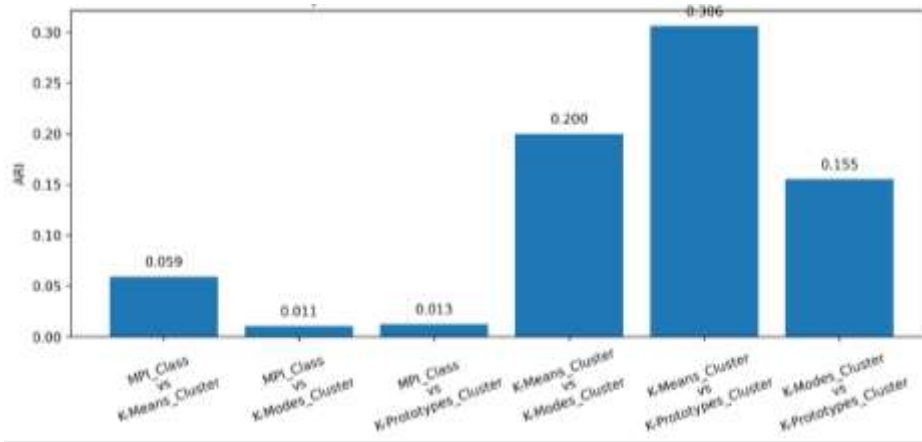


Fig. 5 Adjusted Rand Index plot across framework pair

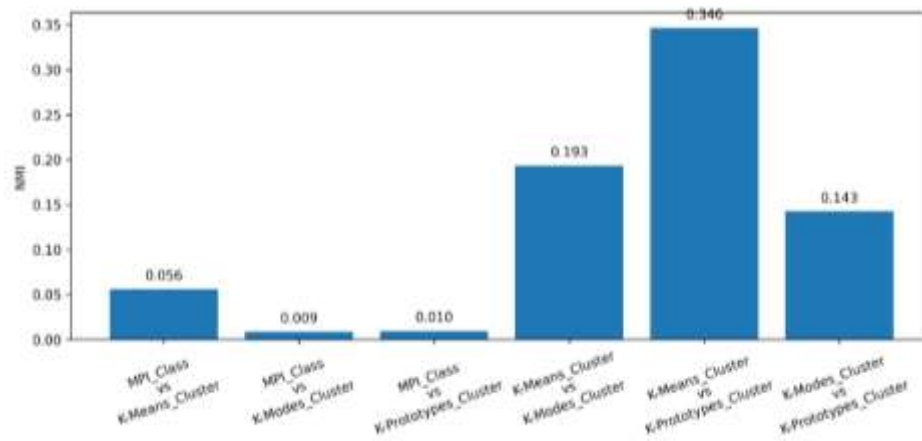
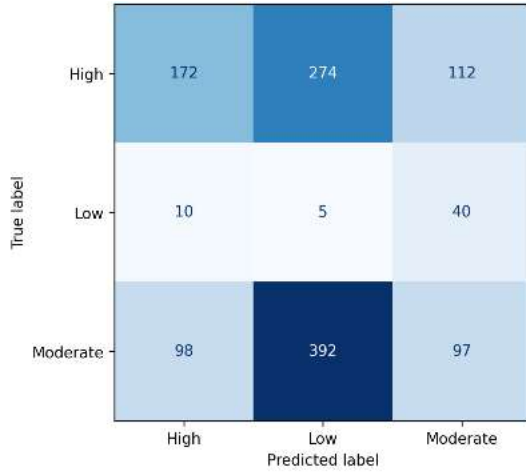


Fig. 6 Normalized Mutual Information plot across framework pair

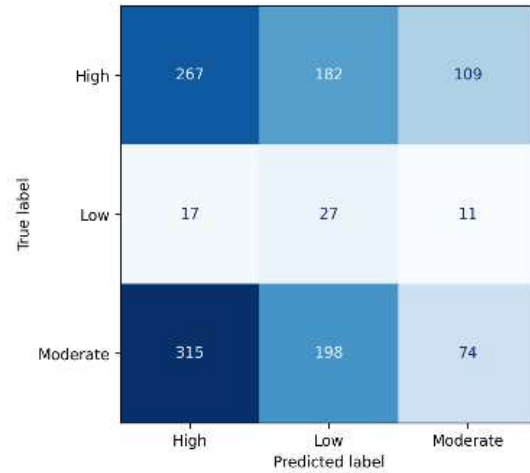
The confusion matrixes of the pairwise analysis are presented on Fig.7. The matrix of MPI with the K-Means clustering (7a) reveals a substantial divergence in household vulnerability assignments. Of the 558 households classified as Highly Vulnerable by MPI, only 172 were similarly identified by K-Means, while 274 were reassigned to the Low vulnerability cluster and 112 to the Moderate cluster. Likewise, the 55 households categorized as Low vulnerability by MPI were dispersed across all three K-Means clusters, with only 5 retaining their original classification, whereas 10 and 40 were reassigned to the High and Moderate clusters, respectively. For the 587 households classified as Moderately Vulnerable by MPI, only 97 remained in the corresponding K-Means category, while 392 were reassigned to the Low cluster and 98 to the High cluster. The limited concentration of observations along the principal diagonal of the confusion matrix, coupled with the widespread off-diagonal allocations, indicates weak correspondence between the two classification frameworks. Specifically, K-Means exhibits a tendency to assign a large proportion of households to the Low vulnerability cluster, whereas the MPI framework classifies many of these same households as High or Moderate vulnerability. These findings further suggest that the two approaches capture different underlying structures of household vulnerability: MPI reflects deprivation based on predefined multidimensional poverty indicators and thresholds, while K-Means uncovers latent groupings driven by overall similarity across the multidimensional feature space. Consequently, the confusion matrix reinforces the agreement analyses by demonstrating that data-driven clustering identifies alternative vulnerability patterns that are not fully represented by the conventional MPI classification. The confusion matrix comparing the MPI with the K-Modes clustering (7b) similarly reveals considerable disagreement in household vulnerability assignments. Of the 558 households classified as Highly Vulnerable by MPI, 267 were consistently identified as High by K-Modes, while 182 and 109 were reassigned to the Low and Moderate vulnerability clusters, respectively. Among the 55 households categorized as Low vulnerability by MPI, only 27 retained the same classification, whereas 17 were reassigned to the High cluster and 11 to the Moderate cluster. Likewise, of the 587 households classified as Moderately Vulnerable by MPI, only 74 remained in the Moderate category under K-Modes, while the majority were redistributed to the High (315) and Low (198) clusters. Although K-Modes exhibits a relatively stronger correspondence with MPI in identifying households with High vulnerability than the K-Means model, the widespread off-diagonal allocations indicate that substantial inconsistencies remain across the two frameworks. These discrepancies suggest that K-Modes, which groups households based primarily on similarities in categorical attributes, uncovers latent vulnerability structures that differ from the threshold-based patterns represented by the MPI. The matrix likewise reinforces the earlier agreement analyses, indicating that the categorical clustering process captures alternative household vulnerability profiles that only partially align with the conventional multidimensional poverty classification.

Comparison of the MPI with the K-Prototypes clustering (7c) indicates a moderate level of correspondence while still revealing substantial differences in vulnerability assignments. Of the 558 households classified as Highly Vulnerable by MPI, only 150 were similarly classified as High by K-Prototypes, whereas 207 and 201 were reassigned to the Low and Moderate vulnerability clusters, respectively. Among the 55 households identified as Low vulnerability by MPI, 17 retained the same classification, while 16 and 22 were reassigned to the High and Moderate clusters. For the 587 households categorized as Moderately Vulnerable by MPI, 283 remained within the Moderate cluster under K-Prototypes, while 106 and 198 were reassigned to the High and Low clusters, respectively. Compared with the K-Means and K-Modes results, K-Prototypes demonstrates a more balanced distribution of households across the three vulnerability strata,

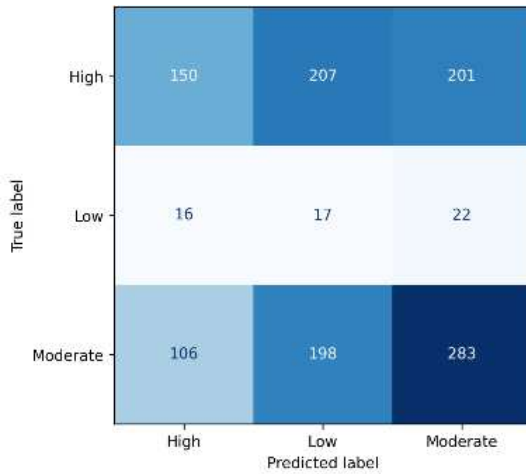
particularly in preserving a larger proportion of the Moderate vulnerability group. Nevertheless, the relatively large number of off-diagonal classifications indicates that notable discrepancies still exist between the MPI framework and the mixed-data clustering approach. These findings suggest that by simultaneously incorporating both numerical and categorical household characteristics, K-Prototypes captures more diverse socioeconomic relationships than the threshold-based MPI classification, while maintaining comparatively greater consistency with the multidimensional poverty framework. The confusion matrix comparing the K-Means and K-Mode's clustering algorithms (7d) reveals substantial differences in how the two unsupervised learning methods partition household vulnerability. Of the 280 households classified as High vulnerability by K-Means, only 46 were similarly identified by K-Modes, while 95 and 139 were reassigned to the Low and Moderate clusters, respectively. The greatest disparity is observed among the 671 households classified as Low vulnerability by K-Means, where 441 were reassigned to the High cluster and 230 remained in the Low cluster under K-Modes, with none assigned to the Moderate cluster. Similarly, of the 249 households categorized as Moderately Vulnerable by K-Means, only 55 retained the same classification, while 112 and 82 were reassigned to the High. Fig.7(e) matrix compares the K-Means and K-Prototypes clustering algorithms, demonstrating a relatively stronger level of correspondence than the other clustering comparisons, although notable differences remain. Out of the 280 households classified as High vulnerability by K-Means, 237 were consistently assigned to the High cluster by K-Prototypes, while 43 were reassigned to the Low cluster and none to the Moderate cluster. Of the 671 households categorized as Low vulnerability by K-Means, only 247 retained the same classification, whereas 408 were reassigned to the Moderate cluster and 16 to the High cluster under K-Prototypes. Similarly, of the 249 households classified as Moderately Vulnerable by K-Means, 98 remained in the Moderate cluster, while 132 and 19 were reassigned to the Low and High clusters, respectively. The relatively high agreement for the High vulnerability category, coupled with the moderate preservation of the Low and Moderate groups, suggests that K-Means and K-Prototypes identify broadly similar vulnerability structures, particularly for the most vulnerable households. However, the considerable redistribution of households between the Low and Moderate clusters indicates differences in how the two algorithms interpret transitional vulnerability states. This pattern is consistent with the earlier Cohen's Kappa, ARI, and NMI results, which identified K-Means and K-Prototypes as the most comparable pair of clustering methods. The findings imply that while both algorithms capture similar underlying household vulnerability patterns, the incorporation of mixed numerical and categorical attributes in K-Prototypes enables a more nuanced differentiation of households with intermediate levels of livelihood vulnerability. Comparing the K-Modes and K-Prototypes clustering algorithms indicates weak correspondence. Of the households classified as High vulnerability by K-Modes, only 57 retained the same classification under K-Prototypes, while the majority were reassigned to the Moderate (362) and Low (180) clusters. Similarly, the Low vulnerability group showed only moderate consistency, with 207 households remaining in the Low cluster, whereas 78 and 122 were reassigned to the High and Moderate clusters, respectively. The Moderate vulnerability group exhibited the weakest agreement, with only 22 households consistently classified, while most were redistributed to the High (137) and Low (35) clusters. The widespread off-diagonal allocations suggest that K-Modes and K-Prototypes identify markedly different household vulnerability structures, reflecting the influence of their underlying clustering mechanisms. This finding aligns with the relatively low Cohen's Kappa, ARI, and NMI values, confirming limited agreement between the two methods.



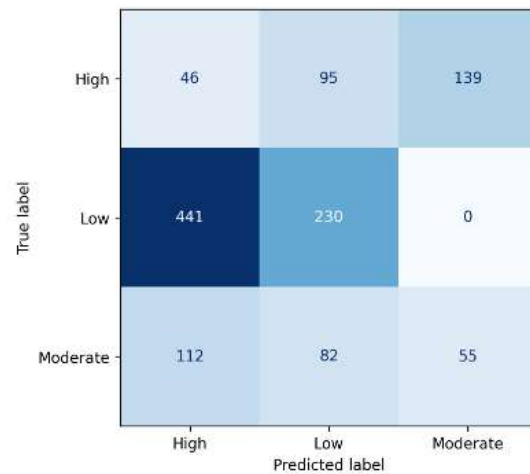
(a) Matrix of MPI vs K-Means



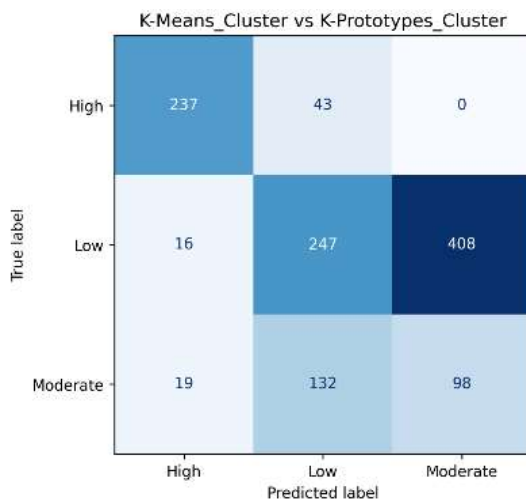
(b) Matrix of MPI vs K-Modes



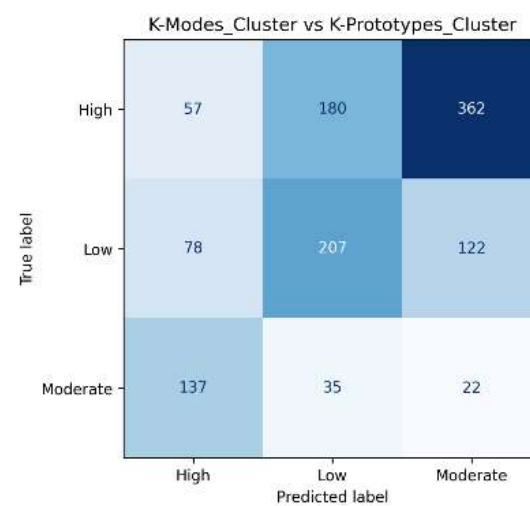
(c) Matrix of MPI vs K-Prototypes



(d) Matrix of K-Means vs K-Modes



(e) Matrix of K-Means vs K-Prototypes



(f) Matrix of K-Modes vs K-Prototypes

Fig. 7 Pairwise confusion matrix of analytical framework per vulnerability classification

Evaluating the level of agreement, similarity, and classification consistency between the frameworks to achieve Objective 2, demonstrates that the four classification frameworks exhibited limited convergence in identifying household livelihood vulnerability, indicating that each captures distinct aspects of the multidimensional feature space. Across the Cohen's Kappa, ARI, and NMI analyses, the MPI framework showed weak agreement with all three clustering methods, suggesting that the threshold-based multidimensional poverty assessment differs substantially from the latent vulnerability structures. Among the intelligent frameworks, K-Means and K-Prototypes consistently exhibited the highest agreement, similarity, and shared information, indicating relatively greater cluster stability and consistency in household partitioning. In contrast, comparisons involving K-Modes generally produced weaker agreement, reflecting its sensitivity to categorical attribute similarities and the resulting divergence from the other methods. The confusion matrixes further revealed extensive household reassignment, with large proportions of households transitioning between High, Moderate, and Low vulnerability categories across frameworks, particularly between the MPI and unsupervised models. These widespread reassignments confirm that vulnerability in Ode-Omi community is neither discrete nor rigid but rather exists along a continuum shaped by complex socioeconomic and infrastructural interactions.

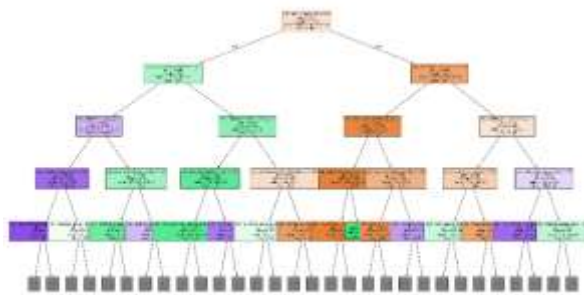
4.4 Explainability of complementary dimensions and domain-invariant determinants of poverty stratification

Objectives 3-4 are achieved in this section, adopting the Decision Tree XAI functionality to examine how service-access, continuous, categorical, and mixed-data representations capture complementary dimensions of household vulnerability. Domain-invariant determinants of vulnerability were also identified as consistent and dominant vulnerability drivers across the analytical frameworks. The Decision Tree modelling demonstrated clear domain-invariant household characteristics across vulnerability stratification and framework variation. The model achieved the best performance on all K-Prototypes metrics based on 10-fold cross-validation (Table 4). These indicate that its vulnerability classifications are the most internally consistent and are best explained by the underlying socioeconomic and infrastructural attributes. This outcome corroborates the earlier agreement analysis, where K-Prototypes exhibited the closest correspondence with the other analytical frameworks. The decision tree of K-Modes on Fig.8(a) and extracted rule on Table 5 returns food security as the principal gateway into household vulnerability, with meal skipping serving as the root determinant of all subsequent classifications. Once food sufficiency is established, the model progressively evaluates access to clean water, healthcare affordability, employment, climate exposure, financial inclusion, and essential infrastructure, indicating that livelihood vulnerability is fundamentally multidimensional, based on the categorical representations. Rather than relying on demographic characteristics alone, the extracted rules reveal that vulnerability is shaped by interacting deficits in food, water, health, infrastructure, and economic opportunity. The K-Means decision tree Fig.8(b) and extracted rules respectively prove that the clusters are primarily differentiated by healthcare accessibility, internet connectivity, electricity and water infrastructure, food security, and financial inclusion. Economic variables such as household income and education further refine the classification, while climate exposure and social inclusion indicators (women's decision-making, caregiver status, discrimination, and community violence) provide additional explanatory power. On Fig.8(c), K-Prototypes decision tree rules show that household clusters are primarily distinguished by access to essential services, particularly water, healthcare, internet, and electricity, before considering food security and economic status. Once these structural conditions are established, income,

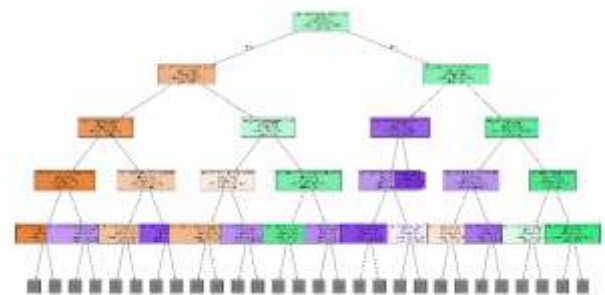
employment, education, household composition, and women's decision-making further refine cluster membership. The MPI decision tree on Fig.8(d) establishes that multidimensional poverty is driven by the interaction of economic, social, infrastructural, health, educational, and environmental factors rather than by income alone, as validated by the extracted rule. The placement of farmland access at the root node identifies productive asset ownership as the strongest determinant of poverty status, while repeated occurrences of healthcare access, food security, water access, education, financial inclusion, and internet connectivity indicate that these dimensions collectively shape household well-being. Furthermore, social variables such as women's decision-making, discrimination, school enrolment, and community violence reveal that social inclusion is equally important in determining poverty outcomes. Tables 5-7 collectively demonstrate that the four analytical frameworks capture complementary yet highly consistent dimensions of household vulnerability despite operating on different data representations. The resulting vulnerability profiles disclose strong convergence in the characteristics defining high, moderate, and low vulnerability, particularly regarding healthcare access, food security, income, education, clean water, employment, and women's participation in household decision-making. Synthesizing these findings, Table 6 and Fig.9 translates modelled heterogeneous vulnerability determinants into a coherent SDG-oriented evidence base for improving the Ode-Omi QoL. The SDG mapping is broadest for the mixed-data representation, spanning 11 SDGs, reflecting its capacity to integrate infrastructure, socioeconomic, demographic, and governance dimensions, while the continuous and categorical representations capture complementary material and social dimensions. Across the SDG frequency distribution, SDG 3 (Good Health and Well-being) emerges most prominently with 5 occurrences, followed by SDGs 1, 6 and 8 (4 occurrences each), indicating that health, basic services, poverty reduction and livelihoods constitute the most persistent policy priorities. Conversely, SDGs 2, 11 and 16 appear less frequently, suggesting more targeted rather than pervasive intervention requirements.

Table 4 Comparative performance of decision tree explainability across the four household vulnerability classification frameworks

Classification Framework	Accuracy	Precision	Recall	F1-Measure	Cohen's Kappa	ROC-AUC
MPI	0.7783	0.7747	0.7783	0.7761	0.5878	0.7970
K-Means	0.8508	0.8510	0.8508	0.8509	0.7471	0.8778
K-Modes	0.7650	0.7655	0.7650	0.7651	0.6146	0.8105
K-Prototypes	0.8717	0.8723	0.8717	0.8719	0.8017	0.9003



(a) Multidimensional vulnerability decision tree of K-Modes



(b) Multidimensional vulnerability decision tree of K-Means

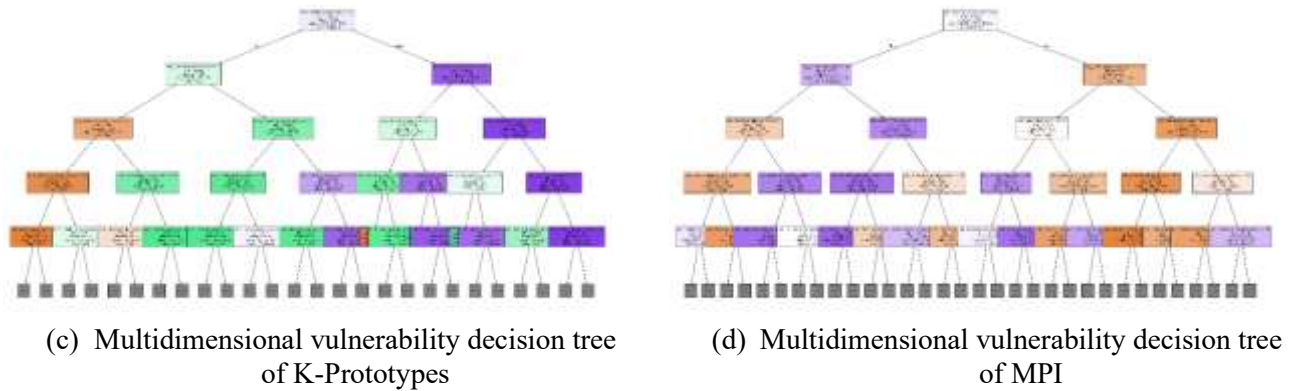


Fig. 8 Decision Trees of the analytical frameworks

Table 5 Brief comparative decision rules for household vulnerability stratification across frameworks as extracted from respective decision rules

Framework	Decision Rule
K-Modes	IF <i>clean_water_access</i> = No AND <i>employment_status</i> = No AND <i>food_security_score</i> = Low AND <i>healthcare_affordability</i> ≠ Sometimes, THEN Cluster 2 (High Vulnerability); otherwise classify according to subsequent branches.
K-Means	IF <i>health_access_score</i> ≤ 5.5, classify households primarily using internet access, electricity access, cooking fuel, water access, food security, income, education, and livelihood characteristics; ELSE (<i>health_access_score</i> > 5.5), classify primarily using women's decision-making, financial inclusion, immunization status, affordable healthcare, climate risk, school enrollment, community violence, clean water access, household size, and socio-economic indicators to assign Cluster 0, 1, or 2.
K-Prototypes	IF <i>water_access_score</i> ≤ 7.5 AND <i>health_access_score</i> ≤ 5.5 AND <i>internet_access_score</i> ≤ 5.5, THEN vulnerability classification is primarily determined by <i>electricity_access_score</i> , with age, monthly income, women's decision-making, meals per day, and food security score serving as secondary discriminating factors.
MPI	IF <i>farmland_access</i> ≠ Yes AND discrimination = Yes AND <i>school_enrollment</i> = Yes AND <i>women_decision_making</i> ≠ Yes AND <i>Health_Acc_Sco</i> ≤ 6.5 AND <i>community_violence</i> = Yes AND <i>income</i> ≤ ₦10,510, THEN the household is classified as Class 2 (Very High Multidimensional Deprivation); otherwise, subsequent decisions are determined primarily by <i>Health_Acc_Sco</i> , healthcare affordability, internet access score, food security score, water access score, and income, with education, employment, household size, clean water access, climate risk, monthly income, and other socioeconomic indicators serving as secondary discriminating factors.

Table 6 Comparative vulnerability profiles revealing domain-invariant household characteristics across vulnerability strata

Framework	High Vulnerability	Moderate Vulnerability	Low Vulnerability
K-Modes	Poor social conditions characterized by lack of clean water, poor sanitation, unaffordable healthcare, limited financial inclusion, unemployment, food insecurity, discrimination, and weak women's participation in household decision-making.	Households exhibiting mixed social conditions, with some access to essential services but deficiencies in healthcare affordability, sanitation, financial access, or food security.	Households with clean water, improved sanitation, affordable healthcare, financial inclusion, school enrolment, employment, and stronger social inclusion indicators.
K-Means	Low scores in water, electricity, internet, food security, and health access, typically accompanied by fewer meals per day and lower household income.	Moderate access scores across infrastructure and basic services, with partial deficiencies in internet, electricity, food security, or healthcare.	High scores in water, electricity, internet, food security, and health access, generally supported by better income levels and adequate daily meals.
K-Prototypes	Low electricity, internet, water, food security, and healthcare access scores combined with lower educational attainment and low income, with age and women's decision-making acting as secondary discriminating factors.	Moderate infrastructure and healthcare access with intermediate food security and education levels; classification further influenced by income and household demographic characteristics.	High electricity, internet, water, food security, and healthcare access scores together with higher education and income levels, indicating better overall socioeconomic conditions.
MPI	No farmland access, presence of discrimination, poor school enrolment, weak women's decision-making, poor health access, community violence, and extremely low household income, with healthcare affordability, food security, water access, and internet access further increasing deprivation.	Households experiencing partial deprivation in healthcare, education, food security, livelihood resources, or infrastructure, but not across all MPI dimensions simultaneously.	Farmland access, affordable healthcare, school enrolment, financial inclusion, clean water, immunization, adequate income, and favourable health and education conditions, indicating low multidimensional deprivation.

Table 7 Distribution of domain-invariant determinants of household vulnerability per SDG requirements

Domain Representation	Dominant Vulnerability Determinants	SDG Requirements	Evidence Across Frameworks	Domain-Invariant Insight
Service-access representation	Health access, clean water, electricity access, internet access, sanitation, affordable healthcare	SDG 3, SDG 6, SDG 7, SDG 9, SDG 11	Consistently dominant in K-Means/K-Prototypes/MPI	Access to essential services constitutes the most stable predictor of household vulnerability irrespective of data representation.
Continuous representation	Health access score, water access score, electricity access score, internet access score, food security score, household income	SDG 1, SDG 2, SDG 3, SDG 6, SDG 7, SDG 8, SDG 9	Primary splitting variables in K-Means/K-Prototypes/MPI	Quantitative access indicators provide the strongest discriminatory power for separating vulnerability strata.
Categorical representation	Employment status, school enrolment, women's decision-making, discrimination, financial inclusion, farmland access, community violence, immunization, chronic illness	SDG 1, SDG 2, SDG 3, SDG 4, SDG 5, SDG 8, SDG 10, SDG 16	Core decision variables in K-Modes/MPI, with supporting influence in K-Means/K-Prototypes	Socio-demographic and governance-related conditions consistently explain household resilience beyond infrastructure access alone.
Mixed-data representation	Combined infrastructure access, socioeconomic status, demographic characteristics, and governance indicators	SDG 1, SDG 3, SDG 4, SDG 5, SDG 6, SDG 7, SDG 8, SDG 9, SDG 10, SDG 11, SDG 16	Explicitly integrated in K-Prototypes/MPI; reflected through complementary patterns in K-Means/K-Modes	Integrating heterogeneous data captures complementary dimensions of vulnerability and provides the most holistic representation of household deprivation.

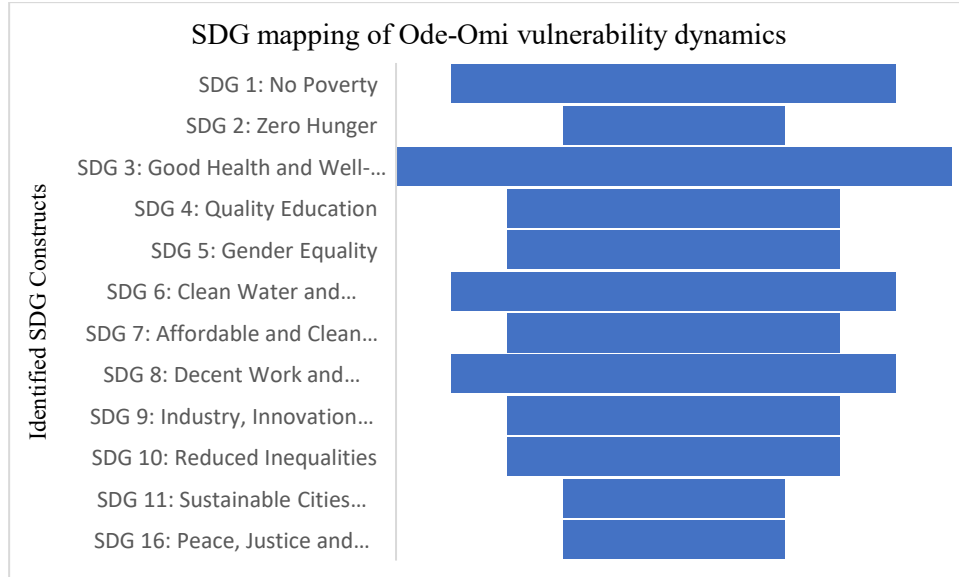


Fig. 9 SDG mapping of domain-invariant Ode-Omi vulnerability dynamics

The radar chart on Fig. 10 reveals clear differences in the extent of feature utilization across the four analytical frameworks. MPI exhibits the broadest feature coverage (27 determinants), and K-Means follows with 20 determinants. K-Prototypes integrates a more selective combination of mixed attributes covering 11 determinants, while K-Modes records the lowest feature utilization (9 determinants). These demonstrate that although the frameworks differ in feature coverage, they consistently identify a common set of domain-invariant vulnerability drivers, highlighting the complementary strengths of heterogeneous data representations for household vulnerability assessment. The ranked horizontal bar chart on Fig.11(a) further reveals the relative importance of individual determinants across the four analytical frameworks. Food Security and Women Inclusiveness emerged as the most consistently influential determinants, each appearing in all four frameworks as evident in the heatmap on Fig.11(b), underscoring their domain-invariant importance in household vulnerability assessment. A second tier of most recurrent determinants included household income, health access, water access, electricity access, internet access, employment status, school enrolment, affordable healthcare, household size, and education. This reflects the complementary roles of socioeconomic capacity and service accessibility. In contrast, age, climate risk, immunization access, source of income, and firewood cooking exhibited comparatively limited influence, while meal skipping, food aid availability, food aid dependency, chronic illness prevalence, and caregiver status of household heads contributed only marginally. Interestingly, while Food Security Score was consistently selected, Nutrition Deprivation Score was excluded by all frameworks, suggesting that the broader food security construct subsumes nutritional deprivation and therefore provides greater discriminatory power for vulnerability stratification. Similarly, Climate Resilience Advantage, Environmental Sanitation Regulations, and Gender were not selected by any framework, indicating minimal contribution to cluster separation within the study population. The heat map further highlights the distinctive behavior of the MPI framework. Unlike the intelligent models, the MPI framework uniquely incorporated meal skipping, food aid dependency, caregiver status of household head, and chronic illness prevalence into its decision structure. Consequently, although all frameworks converge on a common set of

domain-invariant vulnerability determinants, the MPI framework provides the most comprehensive representation of household deprivation by integrating additional welfare-specific indicators that are not exploited by conventional clustering methods. The framework similarity matrix on Fig.12 further demonstrates the complementary relationships among the four analytical approaches. K-Means and MPI exhibited the highest similarity (0.52), indicating that both frameworks rely on a comparable set of dominant vulnerability determinants, particularly service-access and socioeconomic indicators, despite their different analytical foundations. K-Prototypes showed moderate similarity with K-Means (0.41), reflecting its ability to integrate both continuous and categorical information while preserving many of the continuous-data characteristics emphasized by K-Means. K-Modes displayed only weak similarity with K-Means (0.33) because it relies exclusively on categorical attributes. Notably, K-Modes and K-Prototypes exhibited a negative similarity (-0.18), suggesting that the determinants emphasized by purely categorical clustering differ substantially from those prioritized by the mixed-data framework.

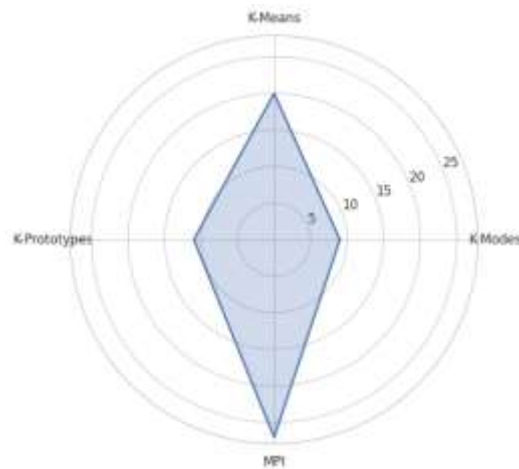


Fig. 10 Frequency of feature occurrence across analytical frameworks

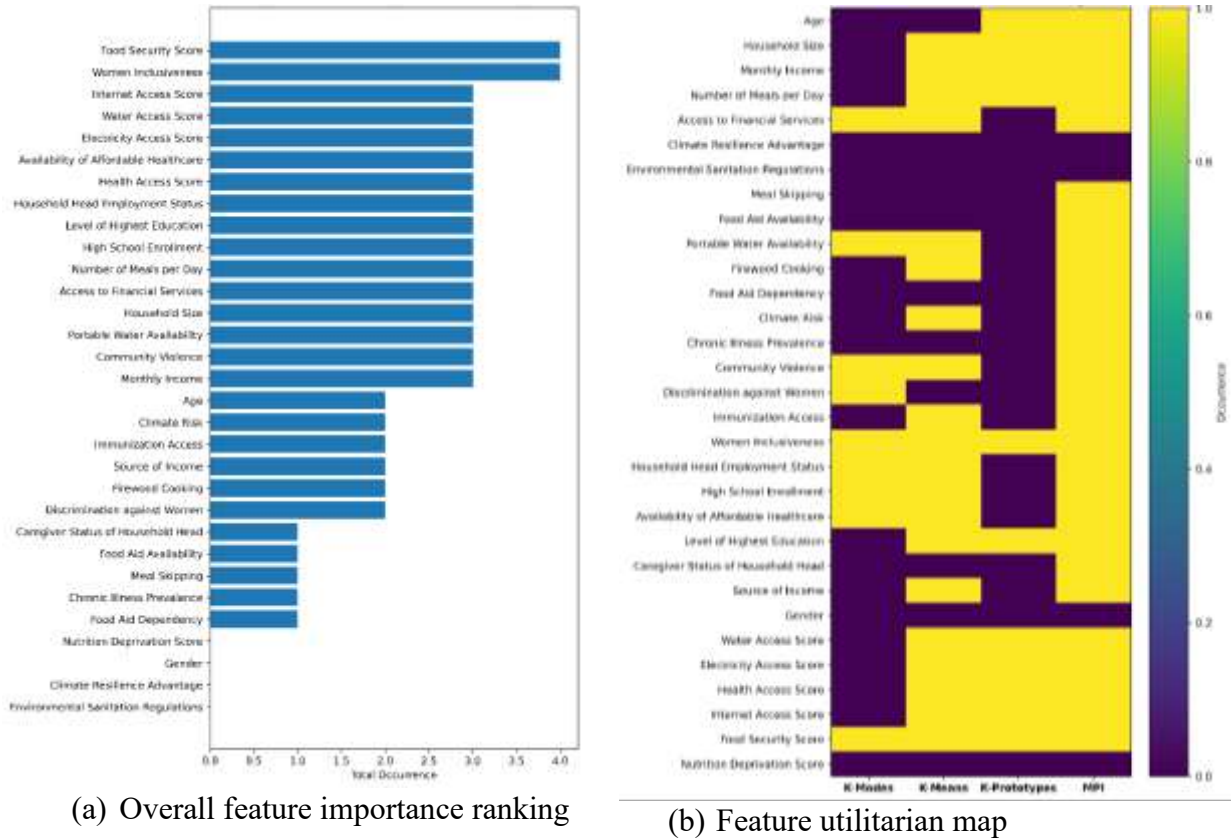


Fig. 11 Feature ranking and utilitarian values across analytical frameworks

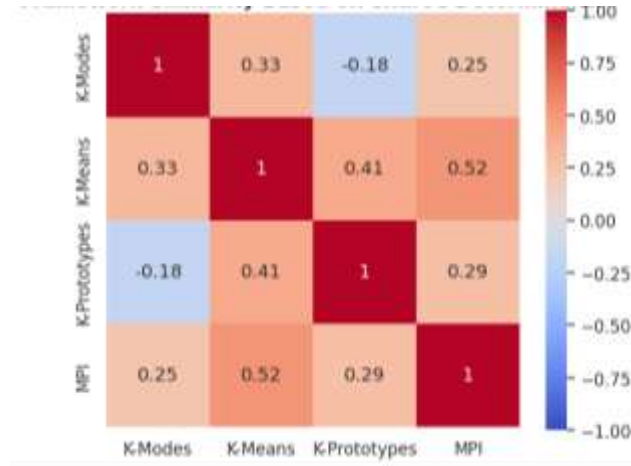


Fig. 12 Framework similarity matrix based on shared determinants

4.5 Discussion of findings

Indications from the experimental results, feature occurrence analysis, similarity matrix, and domain-invariant analysis returns the following policy intelligence findings presented on Table 8.

The findings were aggregated across the experimentation pipelines since none of the four frameworks can be regarded as universally superior because each optimizes a different analytical objective.

4.5.1 Implications for sustainable resource allocation

The findings provide a robust evidence base for precision resource allocation in Nigeria's Delta communities by demonstrating that household vulnerability is multidimensional and cannot be adequately captured using a single analytical framework. The domain-invariant determinants identified across all frameworks offer a stable set of indicators for prioritizing intervention investments. Consequently, resource allocation can transition from conventional uniform distribution toward risk-informed, household-specific targeting, ensuring that public investments are directed to communities exhibiting the greatest composite vulnerability. This evidence directly supports the implementation priorities of Nigeria's 10-Year Agricultural Transformation Plan (2026–2035) (Bahta et al., 2026) and the National Agricultural Investment Plan (Nwozor and Olanrewaju, 2020), both of which emphasize evidence-driven investment planning.

4.5.2 Implications for community resilience building

The complementary nature of the four analytical frameworks demonstrates that resilience extends beyond income poverty to encompass infrastructure accessibility, governance inclusion, healthcare availability, food systems, and social protection. The consistent emergence of women's inclusiveness, food security, health access, clean water, and employment across the analytical frameworks suggests that strengthening these determinants simultaneously is likely to produce the greatest resilience gains. Conversely, framework-specific determinants including discrimination, community violence, farmland access, and digital connectivity indicate that resilience pathways remain context dependent and therefore require localized intervention strategies. These findings support integrated resilience-building programmes that combine agricultural development, healthcare improvement, digital inclusion, social protection, and gender-responsive governance rather than isolated sectoral interventions. Such an integrated perspective aligns closely with the objectives of the African Union Comprehensive Africa Agriculture Development Programme (Africa Union, 2003), which advocates resilient food systems, inclusive agricultural transformation, and sustainable rural livelihoods, while also reinforcing the resilience priorities outlined within Nigeria's agricultural development agenda.

4.5.3 Implications for transformative planning

Beyond household classification, implemented framework provides an operational decision-support mechanism for transformative development planning. The identification of both universal and framework-specific vulnerability determinants enables policymakers to distinguish between long-term structural drivers requiring strategic investment and localized drivers requiring targeted interventions. Rather than relying solely on static poverty indices, planners can integrate multidimensional, categorical, continuous, and mixed-data evidence to anticipate emerging vulnerability patterns and evaluate alternative intervention scenarios. Such an adaptive planning paradigm is particularly relevant for Nigeria's Delta communities, where environmental uncertainty, livelihood diversification, and socioeconomic disparities interact dynamically. By enabling proactive prioritization of infrastructure development, climate adaptation, agricultural investment, healthcare access, and social inclusion programmes, the framework supports evidence-based governance consistent with the objectives of the ECOWAS Environmental Action Plan

(2020–2026) (Tosho, 2025) for sustainable environmental management and climate resilience, while simultaneously complementing Nigeria's long-term agricultural transformation policies through intelligent, data-driven planning capable of supporting sustainable rural development.

4.5.4 Implications for SDGs and QoL Improvement

The study has important implications for achieving the United Nations SDGs by demonstrating that household vulnerability extends beyond income poverty to encompass deficits across other deprivations. The domain-invariant vulnerability determinants aligns directly with SDG 1 (No Poverty), SDG 2 (Zero Hunger), SDG 3 (Good Health and Well-being), SDG 4 (Quality Education), SDG 5 (Gender Equality), SDG 6 (Clean Water and Sanitation), SDG 8 (Decent Work and Economic Growth), and SDG 10 (Reduced Inequalities). Furthermore, the findings provide a multidimensional evidence base for QoL measurement, showing that improvements in essential service access, social inclusion, and livelihood resilience are likely to produce greater welfare gains than income-focused interventions alone. By identifying both universal and context-specific vulnerability drivers, the methodology enables policymakers to monitor household well-being more holistically, prioritize integrated interventions, and develop adaptive policies capable of simultaneously advancing sustainable development, human welfare, and long-term community resilience in Nigeria's delta regions.

Table 8 Summary discussion of findings for transformative planning in Nigeria delta communities

Findings	Policy Intelligence Insight
Convergence of household vulnerability strata	All four analytical frameworks consistently identified three broad household livelihood strata comprising Low (Resilient), Moderate (Transitional), and High Vulnerability households. This convergence demonstrates that vulnerability is not random but naturally stratified irrespective of whether categorical, continuous, mixed-data, or multidimensional poverty representations are employed, thereby providing a stable foundation for policy prioritization.
Complementary vulnerability representations	Although the frameworks converged on the same three strata, they differed substantially in the households assigned to each category and in the variables driving those assignments. The moderate similarity between K-Means and MPI (0.52), together with the low similarity between K-Modes and K-Prototypes (-0.18), indicates that conventional multidimensional poverty measures may overlook complex nonlinear interactions that are captured through ML-based clustering. These supports combining multiple analytical perspectives for evidence-based policy intelligence.
Framework-specific vulnerability pathways	Each framework emphasized distinct determinants of household vulnerability. MPI primarily highlighted structural deprivation through farmland access, discrimination, community violence, school enrolment, health access, and income, whereas K-Means emphasized quantitative service-access indicators such as electricity, water, internet, food security, and healthcare access. K-Modes focused on categorical social exclusion variables including employment, financial inclusion, healthcare affordability, and clean water, while K-Prototypes integrated both infrastructure and socioeconomic characteristics. These differences reveal multiple context-specific pathways into vulnerability requiring tailored intervention strategies.
Emergence of domain-invariant vulnerability determinants	Despite methodological differences, several determinants consistently emerged across all four frameworks, notably food security and women's inclusiveness. Health access, water access, household income, education, employment, clean water access, and affordable healthcare were also prominent. These variables constitute a set of domain-invariant vulnerability determinants, representing robust policy indicators that remain influential for QoL improvement, irrespective of data representation or analytical framework.

Differential policy value of contextual indicators	Variables such as meal skipping, food aid availability, food aid dependency, caregiver status, chronic illness, and community violence were predominantly identified by the MPI framework, while variables including gender, climate resilience advantage, environmental sanitation regulations, and nutrition deprivation score contributed minimally across the analytical frameworks. This finding suggests that certain contextual indicators are highly policy-specific rather than universally discriminative, enabling policymakers to distinguish between indicators suitable for targeted interventions and those appropriate for broad population-level vulnerability monitoring.
--	---

5 Conclusion

This study investigated factors that determined quality of life in Nigeria's delta community of Ode-Omi through domain-invariant analytics of demographic, service-access and categorical vulnerability dynamics of households. The aim was to determine if multidimensional data representations of deprivation dynamics within the region would reveal similar or diverse vulnerability with respect to comparative unsupervised modelling. Decision Tree offered explainability functionalities to cross-check patterns observed from ground truths established by Multidimensional Poverty Index, K-Means, K-Modes and K-Prototype paradigms. Findings indicate that household vulnerability in the Ode-Omi cannot be adequately represented using a single data type, for a reliable, inclusive and robust quality of life assessment. Rather, mixed-data learning provided the most faithful representation of the complex interactions among socioeconomic, infrastructural, demographic, and governance-related determinants, thereby offering the strongest evidence base for vulnerability profiling and policy-targeted interventions. Quality of life determinants emerge as robust, domain-invariant vulnerability drivers that remain consistently influential regardless of whether the framework models categorical, continuous, mixed, or multidimensional poverty representations. Each framework emphasized distinct determinants of household vulnerability. MPI primarily highlighted structural deprivation through farmland access, discrimination, community violence, school enrolment, health access, and income, whereas K-Means emphasized quantitative service-access indicators such as electricity, water, internet, food security, and healthcare access. K-Modes focused on categorical social exclusion variables including employment, financial inclusion, healthcare affordability, and clean water, while K-Prototypes integrated both infrastructure and socioeconomic characteristics. These differences reveal multiple context-specific pathways into vulnerability requiring tailored intervention strategies, as envisaged in the main aim of the study. The findings indicate that no single analytical framework is universally optimal; rather, their complementary integration provides a more comprehensive and explainable characterization of household deprivation than any standalone approach.

Ethic Statement

Participation in the study was voluntary, and informed consent was obtained from all participants. The final section of the survey explicitly sought participants' consent for their responses to be used for research purposes, and all participants indicated their consent. Responses were anonymized and treated confidentially and were used solely for the purposes of this research.

Declarations

Data Availability

The datasets analyzed during the current study are available upon request

Conflict of interest

The authors have no relevant interests to disclose.

References

- Cordes, D., Sefah, P., & Marinova, D. (2024). Developing digital twins of urban low-income communities in Sub-Saharan Africa: a case study in Ghana, West Africa. *Discover Analytics*, 2(1), 18.
- 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.
- Khan, S. U., & Sloboda, B. W. (2023). Spatial analysis of multidimensional poverty in Pakistan: Do income and poverty score of neighboring regions matter?. *GeoJournal*, 88(3), 2823-2849.
- Mathenge, M., Sonneveld, B. G., & Broerse, J. E. (2023). Mapping the spatial dimension of food insecurity using GIS-based indicators: A case of Western Kenya. *Food Security*, 15(1), 243-260.
- Suhardono, S., Fitria, L., Prayogo, W., Lee, C. H., & Suryawan, I. W. K. (2025). Enhancing community engagement with digital twins: Technological adoption in marine debris management. *Journal of Urban Management*.
- Liu, M., Ge, Y., Hu, S., & Hao, H. (2023). The spatial effects of regional poverty: Spatial dependence, spatial heterogeneity and scale effects. *ISPRS International Journal of Geo-Information*, 12(12), 501.
- Ruiu, M. L., & Ragnedda, M. (2024). From poverty to digital poverty. *Digital-Environmental Poverty: Digital and environmental inequalities in the post-covid era*, 15-39.
- Çelik, S., Öztürk, Ö. F., Akkucuk, U., & Şaşmaz, M. Ü. (2025). Global sustainability performance and regional disparities: A machine learning approach based on the 2025 SDG Index. *Sustainability*, 17(16), 7411.
- Ahimah-Agyakwah, S., Nketiah-Amponsah, E., & Agyire-Tettey, F. (2022). Urbanization and poverty in Sub-Saharan Africa: evidence from dynamic panel data analysis of selected urbanizing countries. *Cogent Economics & Finance*, 10(1), 2109282.
- Thorbecke, E. (2023). The interrelationships among growth, inequality and poverty in Africa. *Journal of African Economies*, 32(Supplement_2), ii81-ii86.
- Leichenko, R., & Silva, J. A. (2024). Climate change and poverty. In *Routledge Handbook of Climate Change and Society* (pp. 374-387). Routledge.

- Appiah-Otoo, I., Chen, X., Song, N., & Dumor, K. (2022). Financial development, institutional improvement, poverty reduction: The multiple challenges in West Africa. *Journal of Policy Modeling*, 44(6), 1296-1312.
- Nowak-Garmer, R. (2023). Social protection in Africa. In *Handbook on social protection and social development in the global south* (pp. 112-129). Edward Elgar Publishing.
- Motadi, M. S. (2025). A strategy for poverty eradication towards sustainable development. *OIDA International Journal of Sustainable Development*, 18(09), 141-154.
- Atiglo, D. Y., Abu, M., Jayson-Quashigah, P. N., Addo, K. A., & Codjoe, S. N. A. (2022). Sociodemographic and geophysical determinants of household vulnerability to coastal hazards in the Volta Delta, Ghana. *International Journal of Disaster Risk Reduction*, 78, 103146.
- Hossain, M. S., Ambinakudige, S., Williams, B., Hardy, D., & Mensah, J. (2026). A Place-Based Framework for Assessing Flood-Related Socio-Economic Vulnerability in the Yazoo-Mississippi Delta. *International Journal of Disaster Risk Reduction*, 106338.
- Amoako Johnson, F., Jayson-Quashigah, P. N., Hornby, D., Hill, C., Abu, M., Appeaning Addo, K., & Padmadas, S. (2025). Mapping socio-environmentally vulnerable hotspots in the Volta Delta of Ghana. *Plos one*, 20(5), e0322453.
- Odiana, S., Godwin, O. G., & Okoduwa, A. K. (2026). Integrating Long-Term Climate Trend Analysis and Household Socioeconomic Vulnerability to Assess Climate Stress in the Niger Delta Floodplain (Ughelli, Nigeria). *Trends in Ecological and Indoor Environmental Engineering*, 4(2), 12-21.
- Imide, I. O. (2026). Economic Contribution of Artisanal Fishing Trade to Household Income, Employment, and Local Revenue in Gbaramatu Kingdom, Delta State, Nigeria. *Journal Of Emerging Trends in Management Sciences and Entrepreneurship*, 8(3), 134-147.
- Onyenekwe, C. S., Sarpong, D. B., Egyir, I. S., Opata, P. I., & Oyinbo, O. (2024). A comparative study of farming and fishing households' livelihood vulnerability in the Niger Delta, Nigeria. *Journal of Environmental Planning and Management*, 67(1), 217-241.
- Aransiola, S. A., Zobeashia, S. L. T., Ikhumetse, A. A., Musa, O. I., Abioye, O. P., Ijah, U. J. J., & Maddela, N. R. (2024). Niger Delta mangrove ecosystem: Biodiversity, past and present pollution, threat and mitigation. *Regional Studies in Marine Science*, 75, 103568.
- Okpebenyo, W., Onoh, C., Cornell, C., & Igwe, A. (2023). Revisiting the resource curse in Nigeria: the case of Niger Delta. *KIU Interdisciplinary Journal of Humanities and Social Sciences*, 4(1), 259-276.
- Gbadamosi, F., & Aldstadt, J. (2025). The interplay of oil exploitation, environmental degradation and health in the Niger Delta: A scoping review. *Tropical Medicine & International Health*, 30(5), 351-367.
- Siloko, B. E. (2024). Human security, sustainable livelihoods and development: the case of the Niger Delta region in Nigeria. *Global Discourse*, 14(2-3), 411-432.

- Agbonifo, P. E. (2022). Socio-economic implications of poor environmental management: a framework on the Niger Delta questions. *Environment, Development and Sustainability*, 24(2), 2453-2470.
- Nkem, A. C., Devine, S., Ogaji, D. S., & Topp, S. M. (2024). Economic exclusion and the health and wellbeing impacts of the oil industry in the Niger Delta region: a qualitative study of Ogoni experiences. *International Journal for Equity in Health*, 23(1), 183.
- Chijioke, N. M., Ozobialu, B. C., Peretomode, E., Chibuzo, N. J., & Nwokoro, E. C. (2025). Causes and effects of environmental degradation in the Niger Delta and the mitigation measures that can be implemented. *Saudi Journal of Humanities and Social Sciences*, 10(2), 45-55.
- Ejikeme, J. N. U. (2025). Impact of tourism development on the quality of life of host communities in the Niger Delta region of Nigeria. *Humanities and Social Sciences Communications*, 12(1), 1809.
- Soremi, T. (2020). The implications of oil theft on social and economic development in the Niger Delta. *Global Journal of Social Sciences*, 19, 1-11.
- Egbe, G. (2025). Assessing the impact of socioeconomic disparities on criminal behaviors in oil producing communities in Niger Delta Nigeria. *Journal of Economic Criminology*, 7, 100119.
- Abobo, B. O., Wilson, O. O., Adamu, E. A., & Ufuoma, O. K. (2026). Economic Strain and Kidnapping in Nigeria's Niger Delta. *Ibadan Journal of Sociology*, 17(2), 20-28.
- Ehizoyanyan, O. F., & Oke, C. I. A. (2024). A comparative analysis of the impact of Niger Delta crisis on Edo and Delta states. *Journal of Current Social and Political Issues*, 2(2), 74-87.
- Kakwani, N., & Son, H. H. (2026). Multidimensional poverty: A New perspective on measurement. *The Journal of Development Studies*, 62(3), 445-459.
- Brando, N., & Pitasse Fragoso, K. (2020). Capability deprivation and the relational dimension of poverty: Testing universal multidimensional indexes. In *Dimensions of Poverty: Measurement, Epistemic Injustices, Activism* (pp. 301-319). Cham: Springer International Publishing.
- Xiao, H., Liang, X., Chen, C., & Xie, F. (2022). The impact of multidimensional poverty on rural households' health: From a perspective of social capital and family care. *International Journal of Environmental Research and Public Health*, 19(21), 14590.
- Bonnerjee, A. (2025). Unfolding under-five multidimensional poverty in india: A machine learning approach. *Available at SSRN 5149987*.
- Gunaratne, M. S., Firdaus, R. R., Sihivaitthasan, A. S., & Kamsi, N. S. (2023). The importance of multidimensional vulnerability analysis in sustainable development. *Journal of Sustainability Science and Management*, 18(9), 38-57.
- Rivas Bustos, D. F., Gutierrez, J. A., & Rueda, S. J. (2026). Automated Vulnerability Scanning and Prioritisation for Domestic IoT Devices/Smart Homes: A Theoretical Framework. *Electronics*, 15(2), 466.

- Saha, S., Ruganzu, L., & Kar, S. (2025). Prediction of multidimensional poverty status with machine learning classification at household level: Empirical evidence from Tanzania. *IEEE Access*, 13, 23461-23471.
- Arranhado, E., Barbosa, L., & Bastos, J. A. (2025). Multidimensional poverty in Benin. *Socio-Economic Planning Sciences*, 99, 102181.
- Olaleye, T., Arogundade, O., Okewale, O., JohnBosco, A., Akintunde, S., Benedict, S., & Makinde, P. (2025). Interrogating Ground Truth Assumptions for Human Factor Software Engineering through Local Interpretable Decision Tree eXplanations.
- Munir, S., Mah, B., Kalsi, K., Kapania, S., Posada, J., Law, E., Wang, D., and Ahmed, S. (2026, June). The Consensus Trap: Dissecting Subjectivity and the “Ground Truth” Illusion in Data Annotation. In *The 2026 ACM Conference on Fairness, Accountability, and Transparency* (pp. 2643-2680).
- Steier, J., Goebel, M., & Iwaszczuk, D. (2024). Is your training data really ground truth? A quality assessment of manual annotation for individual tree crown delineation. *Remote Sensing*, 16(1)
- Agnew, C., Eising, C., Denny, P., Scanlan, A., Van De Ven, P., & Grua, E. M. (2023). Quantifying the effects of ground truth annotation quality on object detection and instance segmentation performance. *IEEE Access*, 11, 25174-25188.
- Bahta, S. T., Enahoro, D. K., Wanyoike, F. N., Maruta, A. A., Bamidele, O., Chan, D., & Baltenweck, I. (2026). The Nigeria Livestock Master Plan (N-LMP): Investment Roadmap (2026–2030).
- Nwozor, A., & Olanrewaju, J. S. (2020). The ECOWAS agricultural policy and the quest for food security: Assessing Nigeria’s implementation strategies. *Development Studies Research*, 7(1), 59-71.
- Africa Union. (2003). Comprehensive Africa agriculture development programme. *Midrand, South Africa: NEPAD*.
- Tosho, I. R. (2025). Towards environmental governance: ECOWAS’s commitment to sustainability in West Africa. *Islamic University Journal of Social Sciences*, 4(3), 5292-609.
- Badmus, G. O., Akinyemi, O. D., Gbadebo, A. M., & Oyedepo, J. A. (2020). Hydrochemical analysis of groundwater quality along the coastal aquifers in part of Ogun Waterside, Ogun State, southwestern Nigeria. *Heliyon*, 6(12).
- Badmus, G. O., Akinyemi, O. D., Gbadebo, A. M., Oyedepo, J. A., & Folarin, G. M. (2021). Assessment of seasonal variation of saltwater intrusion using integrated geophysical and hydrochemical methods in some selected part of Ogun Waterside, Southwest, Nigeria. *Environmental Earth Sciences*, 80(3), 99.
- Olaleye, T., Arogundade, O., Adenusi, C., Misra, S., & Bello, A. (2020, December). Evaluation of image filtering parameters for plant biometrics improvement using machine learning. In *International Conference on Soft Computing and its Engineering Applications* (pp. 301-315). Singapore: Springer Singapore.

Shehadeh, A., Alshboul, O., & Tamimi, M. (2025). Leveraging digital twins for community-driven sustainable WEF nexus management. *Sustainable Futures*, 9, 100722.