

A longitudinal multi-proxy geospatial classification of peri-urban transitions across Community Health Units in coastal Kenya

Prissy Makena

Aga Khan University

Felix Oluoch

`oluoch.felix@aku.edu`

Aga Khan University

Rosebella Iseme-Ondiek

Aga Khan University

Fredrick Gudda

Aga Khan University

Alfred Keter

Aga Khan University

Jai Das

Institute of Global Health and Development

Zulfiqar Bhutta

Institute of Global Health and Development

Anthony Ngugi

Aga Khan University

Research Article

Keywords: urbanization, peri-urban, geospatial analysis, nighttime lights, Sentinel-2, WorldPop, Degree of Urbanization, Community Health Units, Kilifi, Kenya

Posted Date: April 1st, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-9231692/v1>

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

[Read Full License](#)

Additional Declarations: No competing interests reported.

Version of Record: A version of this preprint was published at International Journal of Health Geographics on August 15th, 2026. See the published version at <https://doi.org/10.1186/s12942-026-00483-5>.

A longitudinal multi-proxy geospatial classification of peri-urban transitions across Community Health Units in coastal Kenya

Prissy Makena^{1*}, Felix Oluoch^{1,2*}, Rosebella Iseme Ondiek¹, Fredrick Gudda^{1,2}, Alfred Keter¹, Jai Das^{3,4}, Zulfiqar Bhutta^{3,5}, Anthony Ngugi¹

¹ Department of Population Health, Aga Khan University, Nairobi, Kenya

² Institute for Global Health and Development, Aga Khan University, Nairobi, Kenya

³ Institute for Global Health and Development, Aga Khan University, Karachi, Pakistan

⁴ Department of Paediatrics and Child Health, Aga Khan University, Karachi, Pakistan

⁵ Centre for Global Child Health, The Hospital for Sick Children, Toronto, Ontario, Canada

** Joint first authors*

Corresponding authors:

Prissy Makena*

Department of Population Health, Aga Khan University, Nairobi, Kenya

Email: prissy.makena@aku.edu

Dr Felix Oluoch, PhD*

Department of Population Health, Aga Khan University, Nairobi, Kenya

Email: oluoch.felix@aku.edu

Abstract

Background: Urbanization is spatially heterogeneous in many sub-Saharan African settings. This complicates community health planning when rural and peri-urban labels are applied informally or updated inconsistently. We assessed whether multiple publicly available geospatial proxies could characterize settlement dynamics and support Community Health Unit (CHU) level classification within the Kaloleni Rabai Health and Demographic Surveillance System (KRHDSS) in coastal Kenya.

Methods: We conducted a longitudinal ecological analysis across 10 CHUs in Kilifi County with annual observations from 2017 to 2024. Five geospatial proxy indicators of settlement intensity were summarized within CHU boundaries: Visible Infrared Imaging Radiometer Suite (VIIRS) nighttime lights radiance, Sentinel-2 built-up area percentage, WorldPop built-up area percentage, WorldPop population density, and Degree of Urbanization urban percentage. The built-up proxies quantify the extent of built surfaces and do not distinguish construction materials or housing quality. We quantified year-to-year and cumulative percent change, assessed concordance using Pearson correlation coefficients, and derived a consensus classification using k-means clustering applied separately to standardized CHU-level median proxy values, followed by cross-proxy vote scoring.

Results: The analytic dataset comprised 80 CHU-year observations per proxy with no missing raw proxy values. Degree of Urbanization showed a pronounced floor effect with 37 zero observations. Median cumulative percent change from 2017 to 2024 was 62.25% for nighttime lights, 110.74% for Sentinel-2 built-up area, 34.14% for WorldPop built-up area, and 14.67% for population density. Sentinel-2 built-up area showed substantial interannual variability, with 41.0% of valid year-to-year transitions negative. Pearson correlation coefficients between pairs of proxy indicators, using pooled raw values across all 80 CHU-year observations, ranged from 0.710 to 0.955. Cross-proxy vote scoring classified 3 of the 10 CHUs as consensus peri-urban.

Conclusions: A multi-proxy geospatial approach provides a reproducible complementary framework for delineating CHU settlement contexts in transitional settings. It may support periodic reassessment

49 alongside established administrative classifications where settlement patterns are changing. Divergent
50 proxy patterns, including volatility and floor effects, highlight the value of combining complementary
51 products when distinguishing peri-urban from rural contexts for routine health system planning.

52
53 **Keywords:** urbanization; peri-urban; geospatial analysis; nighttime lights; Sentinel-2; WorldPop;
54 Degree of Urbanization; Community Health Units; Kilifi; Kenya

Background

Urbanization is reshaping population distribution and health system demand worldwide. In 2018, 55% of the global population lived in urban areas, and this proportion is projected to rise substantially by 2050 [1]. In sub-Saharan Africa, urbanization is often rapid but spatially uneven, occurring not only through expansion of major cities but also along transport corridors and within small and medium sized towns [2-5]. These patterns create complex peri-urban mosaics that are poorly captured by static administrative classifications. They also influence health systems by changing population density, mobility, service needs, and the coverage requirements and costs of community-based service delivery platforms [6,7].

In Kenya, the Community Health Unit (CHU) is the foundational operational unit for delivering promotive and preventive services under the national community health strategy [8]. Within the Kaloleni Rabai Health and Demographic Surveillance System (KRHDSS) in coastal Kenya, 10 CHUs provide a defined geographic frame for operationalizing spatial stratifications aligned to routine health system units [9]. However, although Kenya has formal administrative and census-based frameworks for classifying urban areas, these may not fully capture CHU-level settlement context in rapidly changing areas. This creates a rationale for complementary spatial approaches to characterize rural, peri-urban, and transitional settings for service planning [10,11]. Recent advances in Earth observation and gridded demographic modelling enable scalable quantification of settlement intensity [12,13]. WorldPop provides annual high-resolution maps of population distribution and supporting covariates that can help characterize small areas where census information may be limited or outdated [14-16]. Complementary built environment metrics include harmonized settlement products integrating Earth observation and population data [12], Sentinel-2 derived land cover products for fine-scale built-up monitoring [17], Visible Infrared Imaging Radiometer Suite (VIIRS) nighttime lights for human activity and electrification intensity [18], and Degree of Urbanization outputs derived from gridded population data [19]. Despite this proliferation of data streams, important gaps persist in public health

81 applications. Urbanization is operationalized differently across products, and individual proxies often
82 diverge in transitional settings where physical expansion, population concentration, and infrastructure
83 development evolve asynchronously [2-5,11,19].

84 This longitudinal ecological study addressed these gaps by integrating five complementary geospatial
85 proxies, namely nighttime lights intensity, Sentinel-2 built-up area, WorldPop built-up area, WorldPop
86 population density, and Degree of Urbanization, to characterize settlement dynamics across 10 CHUs
87 within the KRHDSS from 2017 to 2024 [14,16-19]. Specifically, we aimed to quantify multi-
88 dimensional settlement trajectories, compare agreement across proxies, and derive a consensus CHU
89 classification that could support routine updating of spatial stratification for community health
90 planning.

91 **Methods**

92 *Study design and setting*

93 We conducted a longitudinal ecological analysis to characterize settlement dynamics across CHUs in
94 Kilifi County, Kenya. The analytic unit was the CHU, with annual observations spanning 2017 to 2024.
95 The analysis was conducted within the KRHDSS platform in Kilifi County, coastal Kenya, which has
96 been described in detail elsewhere [9].

97 The KRHDSS covers Kaloleni and Rabai sub-counties and is centred on Mariakani township. The
98 surveillance area is predominantly rural but includes a peri-urban corridor associated with Mariakani
99 and surrounding transport-linked growth. The area is characterized by high poverty, subsistence
100 farming, informal trade, and recurrent seasonal flooding [9].

101 The KRHDSS comprises 10 administratively defined CHUs under Kenya's community health strategy,
102 each typically covering approximately 500 to 1,000 households, and together these units define the
103 operational geography for the present analysis [8]. The 10 CHUs included in this analysis were Buni,

Kibwabwani, Kombeni, Kwale, Mnazimwenga, Mutsengo, Mwele Kisurutini, Tsangatsini, Viragoni, and Vishakani (Figure 1).

Data sources and proxy indicators

We analyzed five complementary geospatial proxy indicators derived from publicly available datasets. All proxy measures were analyzed as continuous variables and summarized within CHU boundaries using zonal statistics aligned to the semantics of each source variable, rather than by imposing predefined categorical cut-offs [14,16-19].

VIIRS nighttime lights were summarized as mean annual radiance within each CHU. Sentinel-2 built-up extent was summarized as the percentage of CHU area classified as built up. WorldPop built-up extent was similarly expressed as the percentage of CHU area classified as built up. WorldPop population density was summarized as mean population density within each CHU. Degree of Urbanization was summarized as the percentage of CHU area classified as urban [14,16-19]. In this manuscript, 'urban percentage' refers only to the source dataset output, whereas 'rural' and 'peri-urban' refer to the final consensus CHU classification derived from cross-proxy clustering and vote scoring.

Spatial processing and harmonization

CHU boundary polygons were processed in Quantum Geographic Information System (QGIS) version 3.28.3 to extract annual zonal summaries from each raster source [20]. Datasets were harmonized by standardizing CHU names, reshaping source files to one record per CHU per year and merging them into a single panel using CHU name and calendar year as composite keys. Quality checks confirmed one-to-one CHU matching across all five proxy sources, complete annual coverage for all CHUs, and explicit distinction between zero values and missing data.

Statistical analysis

Analyses were conducted in Python 3.10. K-means clustering was implemented using scikit-learn [21].

For each proxy, we summarized baseline values in 2017 and endline values in 2024 using medians and interquartile ranges (IQR) across CHUs. We also examined pooled distributions across all 80 CHU-year observations.

For each CHU and proxy, we calculated relative year-to-year percent change, cumulative percent change from 2017 to 2024, and absolute change from 2017 to 2024 using the following formulas:

Relative year-to-year percent change in year t:

$$\Delta t = \frac{(v_t - v_{t-1})}{v_{t-1}} \times 100$$

Cumulative percent change from 2017 to 2024:

$$\Delta_{cum} = \frac{(v_{2024} - v_{2017})}{v_{2017}} \times 100$$

Absolute change from 2017 to 2024:

$$\Delta_{abs} = v_{2024} - v_{2017}$$

Relative year-to-year change was calculated only when the preceding year value was non-zero.

Cumulative percent change was calculated only when the 2017 baseline was non-zero. For cumulative change, we reported the median and IQR across CHUs with non-zero 2017 baselines and estimated 95% confidence intervals using nonparametric bootstrap resampling with 10,000 replications [23,24].

For year-to-year change, we reported pooled medians and IQRs across all valid transitions and derived 95% confidence intervals using block bootstrap resampling at CHU level with 10,000 replications [23,24].

To assess concordance between proxies, we calculated Pearson correlation coefficients with 95% confidence intervals based on Fisher's z transformation. Correlations were examined for three data structures: pooled raw proxy values across all CHU-year observations, year-to-year percent changes across valid transitions, and cumulative percent changes from 2017 to 2024 across CHUs with non-zero baselines.

To derive an exploratory operational CHU classification, we first calculated the median value of each proxy for each CHU across 2017 to 2024 to capture typical settlement intensity over the full observation period. Each proxy was then standardized to z scores to place all measures on a common scale. For each proxy separately, we applied k-means clustering with $k = 2$ to partition CHUs into lower-intensity and higher-intensity groups, using `random_state = 42` and `n_init = 10` [21]. The cluster with the higher mean standardized value was labelled the higher-intensity cluster. For each CHU, we then summed the number of proxies assigning that unit to the higher-intensity cluster, yielding a vote score from 0 to 5. CHUs with scores of 3 to 5 were classified as consensus peri-urban, whereas those with scores of 0 to 2 were classified as consensus rural. This prespecified majority-vote rule was used as a pragmatic operational framework rather than a validated gold standard. The proportion classified as peri-urban was reported with an exact binomial 95% confidence interval [22].

Zero denominators in percentage change calculations were treated as undefined and were not imputed. This affected 9 Sentinel-2 year-to-year transitions, 33 urban-percentage year-to-year transitions, 1 Sentinel-2 cumulative comparison, and 5 urban-percentage cumulative comparisons. Structural zeros in the Degree of Urbanization metric were retained because they reflect the classification properties of that product rather than missing measurement.

Ethical considerations

This study used publicly available, anonymized, aggregated geospatial datasets summarized at the CHU level. No individual-level human participant data was accessed or analyzed. The study was deemed exempt from ethical review by the Aga Khan University Institutional Ethics Review Committee (Reference: 2013/IREC-65, versions v4–v13), as it involved secondary analysis of non-identifiable data.

Results

The analytic dataset included 10 CHUs observed annually from 2017 to 2024, contributing 80 CHU-year observations per proxy. After harmonization and merging across sources, there were no missing proxy values in the final CHU-year panel. The study area and CHU boundaries are shown in Figure 1, and annual spatial panels for each proxy are provided in Supplementary Figures S1 to S5.

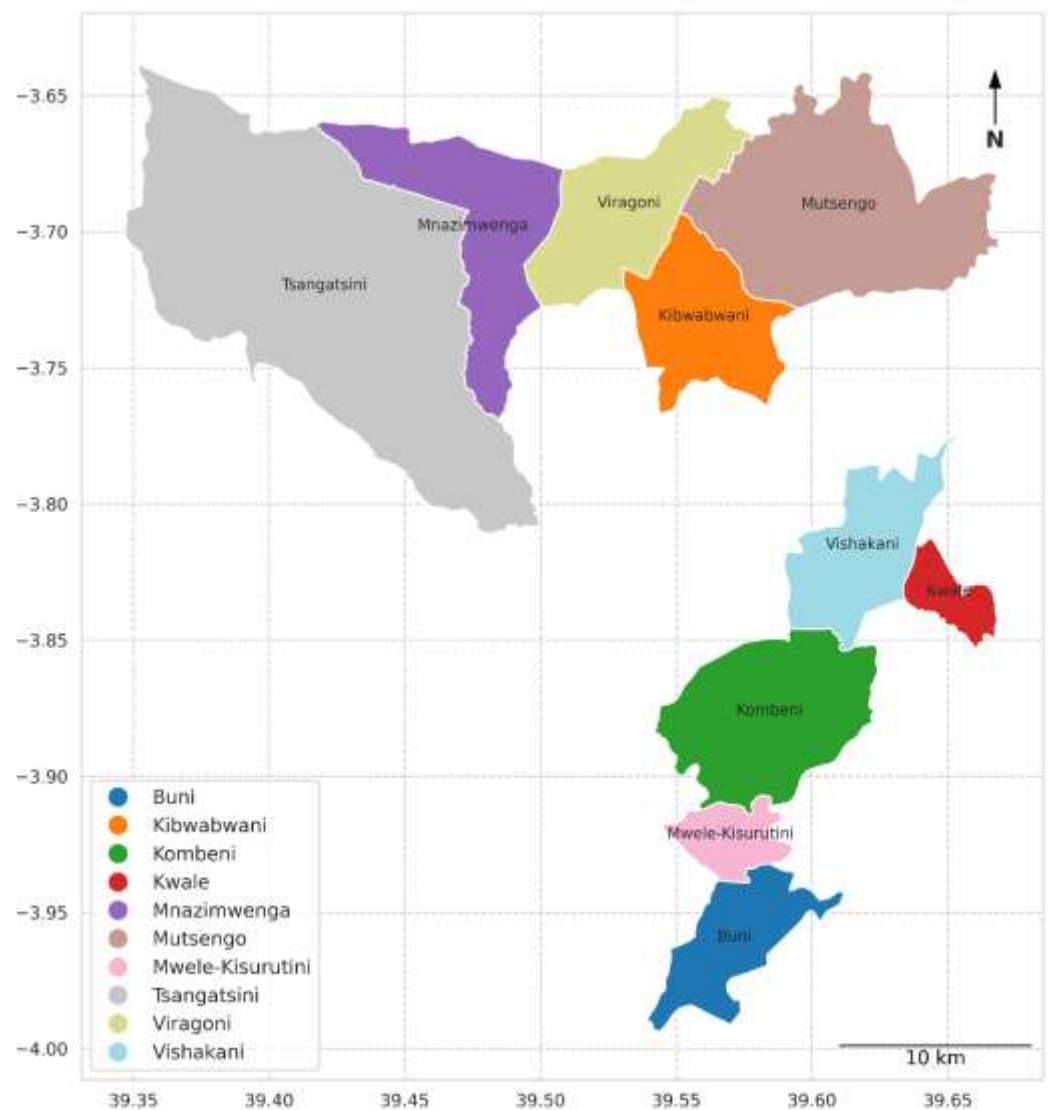


Figure 1. Study area and CHU boundaries in the Kaloleni-Rabai Demographic and Health Surveillance System, Kilifi County, Kenya.

Notes: Polygons represent the 10 CHUs used in all analyses. The map provides geographic context for temporal trends, correlations, and clustering results.

Baseline and endline distributions increased across all five proxies between 2017 and 2024, as summarized in Table 1, while yearly trajectories are shown in Figure 2. Nighttime lights intensity increased from a median of 0.346 (IQR 0.257, 0.523) in 2017 to 0.575 (IQR 0.421, 0.806) in 2024. WorldPop population density increased from 284.6 persons per km² (IQR 110.7, 700.6) to 332.1 persons per km² (IQR 125.0, 807.7). WorldPop built-up area increased from 1.322% (IQR 0.585%, 2.894%) to 2.002% (IQR 0.816%, 3.604%). Sentinel-2 built-up area increased from 3.162% (IQR 0.030%, 18.377%) to 6.419% (IQR 0.349%, 26.851%). Degree of Urbanization urban percentage increased from 26.34% (IQR 0.00%, 65.34%) to 30.40% (IQR 0.00%, 72.57%).

Table 1. Baseline and endline distributions of five urbanization proxies across 10 CHUs in 2017 and 2024.

Proxy	2017 median (IQR)	2024 median (IQR)
Nighttime lights intensity (radiance)	0.346 (0.257, 0.523)	0.575 (0.421, 0.806)
Sentinel-2 built up (%)	3.162 (0.030, 18.377)	6.419 (0.349, 26.851)
WorldPop built-up area (%)	1.322 (0.585, 2.894)	2.002 (0.816, 3.604)
WorldPop population density (persons per km ²)	284.6 (110.7, 700.6)	332.1 (125.0, 807.7)
Degree of Urbanization, urban percentage (%)	26.34 (0.00, 65.34)	30.40 (0.00, 72.57)

Notes: Proxy values are medians and IQRs across 10 CHUs. Nighttime lights intensity values are mean annual radiance. Built-up area values are percentages of CHU area. Population density is persons per km². Urban percentage is from the Degree of Urbanization source dataset.

Degree of Urbanization exhibited a pronounced floor effect, with 37 of 80 observations equal to zero. These zeros were concentrated in four CHUs that remained at 0% urban area across all eight years and one CHU that remained at 0% from 2017 to 2021 (Supplementary Figure S5). Availability of percent change outcomes reflected these floor effects. All 70 possible year-to-year transitions were defined for nighttime lights, WorldPop built-up area, and WorldPop population density, compared with 61 of 70 for Sentinel-2 built-up area and 37 of 70 for urban percentage (Table 2). For cumulative percent change from 2017 to 2024, values were available for all 10 CHUs for nighttime lights, WorldPop built-up area, and population density, compared with 9 CHUs for Sentinel-2 built-up area, and 5 CHUs for urban percentage (Table 2; Supplementary Table S1e).

Table 2. Availability of percent change outcomes for year-to-year and cumulative change calculations by proxy.

Proxy	Total possible transitions	Defined transitions	Undefined transitions	Reason undefined
Nighttime lights intensity	70	70	0	
Sentinel-2 built up	70	61	9	Baseline year value equals 0 for 9 transitions
WorldPop built-up area	70	70	0	
WorldPop population density	70	70	0	
Urban percentage	70	37	33	Baseline year value equals 0 for 33 transitions

Notes: Year-to-year transitions represent changes from 2017–2018 through 2023–2024, yielding 70 possible transitions across 10 CHUs. Percent change outcomes are undefined when the baseline denominator equals zero. Undefined counts arise primarily from floor effects in the Degree of Urbanization metric.

CHU-specific trajectories for all proxies are presented in Figure 2. Across the study period, three broad patterns were apparent: consistently higher intensity profiles in Buni, Mwele Kisurutini, and Vishakani; intermediate profiles in Kombeni and Kwale; and persistently lower intensity profiles in Mnazimwenga, Mutsengo, Viragoni, Kibwabwani, and Tsangatsini (Figure 2). The annual map panels in Supplementary Figures S1 to S5 provide spatial context for these trajectories.

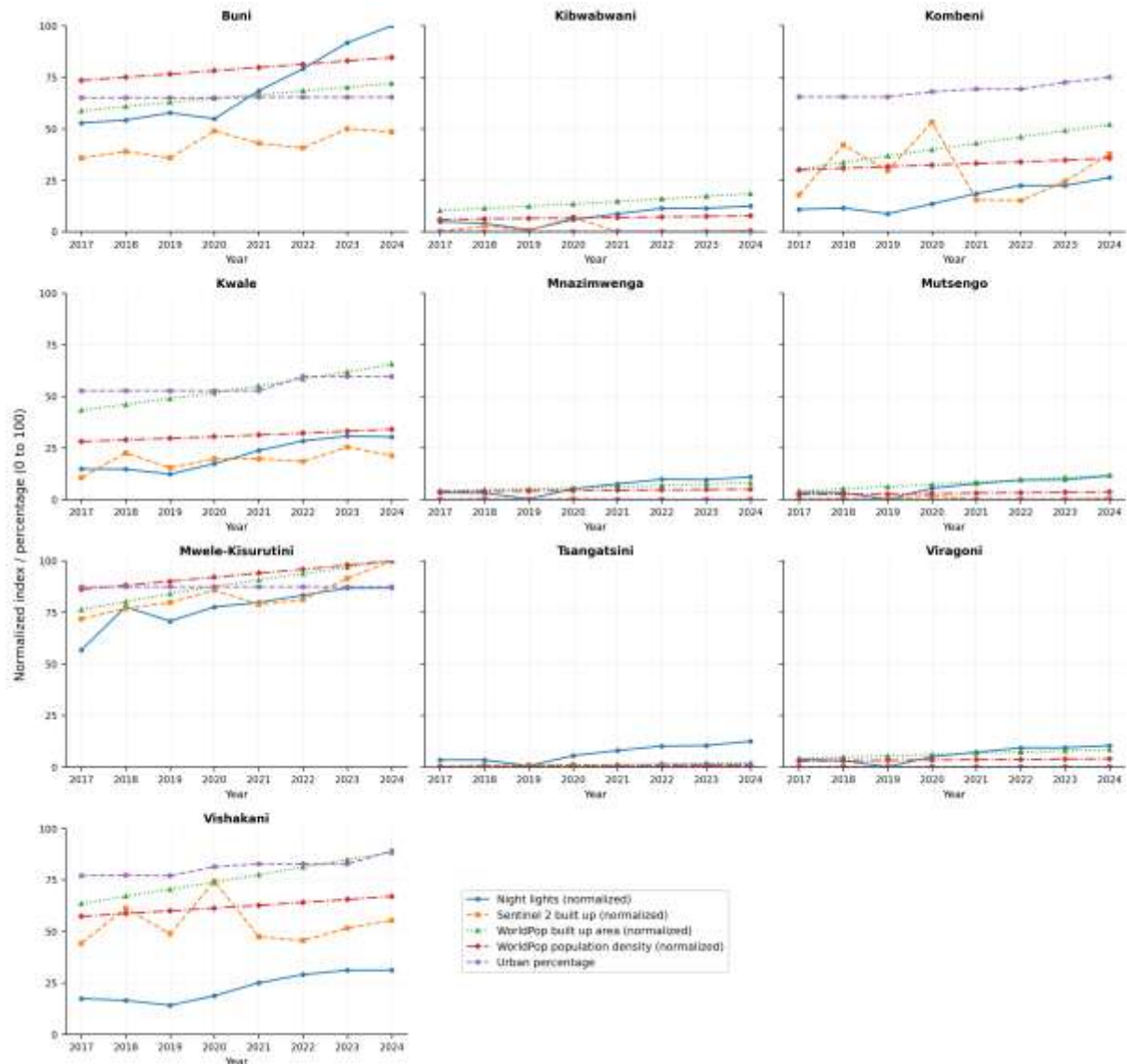


Figure 2. CHU-specific trajectories of five urbanization proxies from 2017 to 2024.

Notes: Lines show annual values for each CHU for nighttime lights intensity, Sentinel-2 built-up percentage, WorldPop built-up percentage, WorldPop population density, and Degree of Urbanization (urban percentage). The figure supports visual comparison of relative intensity and growth patterns across CHUs.

For the primary trend outcome, cumulative percent change from 2017 to 2024, nighttime lights showed a median increase of 62.25% (IQR 51.77%, 69.23%; Supplementary Table S3a). Sentinel-2 built-up area showed the largest median increase, 110.74% (IQR 39.37%, 218.86%; Supplementary Table S3b), although the wide interval reflected small denominators in some CHUs. WorldPop built-up area increased by 34.14% (IQR 28.17%, 43.11%; Supplementary Table S3c), and WorldPop population density increased by 14.67% (IQR 13.59%, 15.42%; Supplementary Table S3d). Among CHUs with

nonzero 2017 baselines, Degree of Urbanization urban percentage increased by 13.15% (IQR 0.48%, 14.57%; Supplementary Table S3e). CHU-specific baseline values, endline values, absolute changes, and cumulative percent changes are presented in Supplementary Tables S1a to S1e.

For year-to-year percent change across available transitions, nighttime lights showed a pooled median increase of 5.90% (IQR -0.20%, 14.85%; Table 3). Sentinel-2 built-up area had a pooled median of 7.21% (IQR -20.33%, 61.18%; Table 3), but 25 of 61 valid transitions (41.0%) were negative despite overall net growth across the full study period (Table 3; Figure 2). WorldPop built-up area showed consistent annual growth with a pooled median of 4.37% (IQR 3.15%, 5.37%; Table 3), and WorldPop population density increased steadily with a pooled median of 2.00% (IQR 1.81%, 2.07%; Table 3). Urban percentage showed minimal year-to-year change, with a pooled median of 0.00% (IQR 0.00%, 0.41%; Table 3).

Table 3. Summary of year-to-year percent change distributions for five urbanization proxies across available transitions.

Proxy	Pooled median (%)	IQR 25th to 75th (%)	Bootstrap 95% CI for median (%)	N transitions	Total possible
Nighttime lights intensity	5.90	-0.20, 14.85	2.38, 11.94	70	70
Sentinel-2 built up	7.21	-20.33, 61.18	-3.08, 28.00	61	70
WorldPop built-up area	4.37	3.15, 5.37	3.99, 4.89	70	70
WorldPop population density	2.00	1.81, 2.07	1.90, 2.03	70	70
Urban percentage	0.00	0.00, 0.41	0.00, 0.00	37	70

Notes: Year-to-year percent change is calculated as (value at year t minus value at year t-1) divided by value at year t-1, multiplied by 100. Transitions are excluded when the prior year value equals zero. Pooled summaries aggregate all defined transitions across CHUs. Negative transitions represent year-to-year decreases.

Correlations between pooled raw proxy values across the 80 CHU-year observations were uniformly strong (Figure 3; Supplementary Table S4A). Pearson correlation coefficients ranged from 0.710 to 0.955, with all p values below 0.0001. WorldPop built-up area and urban percentage showed the strongest association ($r = 0.955$, 95% CI 0.931 to 0.971), while nighttime lights and urban percentage showed the weakest ($r = 0.710$, 95% CI 0.581 to 0.804). Among year-to-year percent changes, only WorldPop built-up area and population density showed a statistically significant correlation ($r = 0.484$,

95% CI 0.281 to 0.646; $p < 0.0001$; Supplementary Table S4B). Correlations of cumulative percent changes were not statistically significant and were imprecisely estimated for proxies with small effective sample sizes, particularly urban percentage and Sentinel-2 built-up area (Supplementary Table S4C). Full correlation matrices with confidence intervals and p values are reported in Supplementary Tables S4A to S4C.

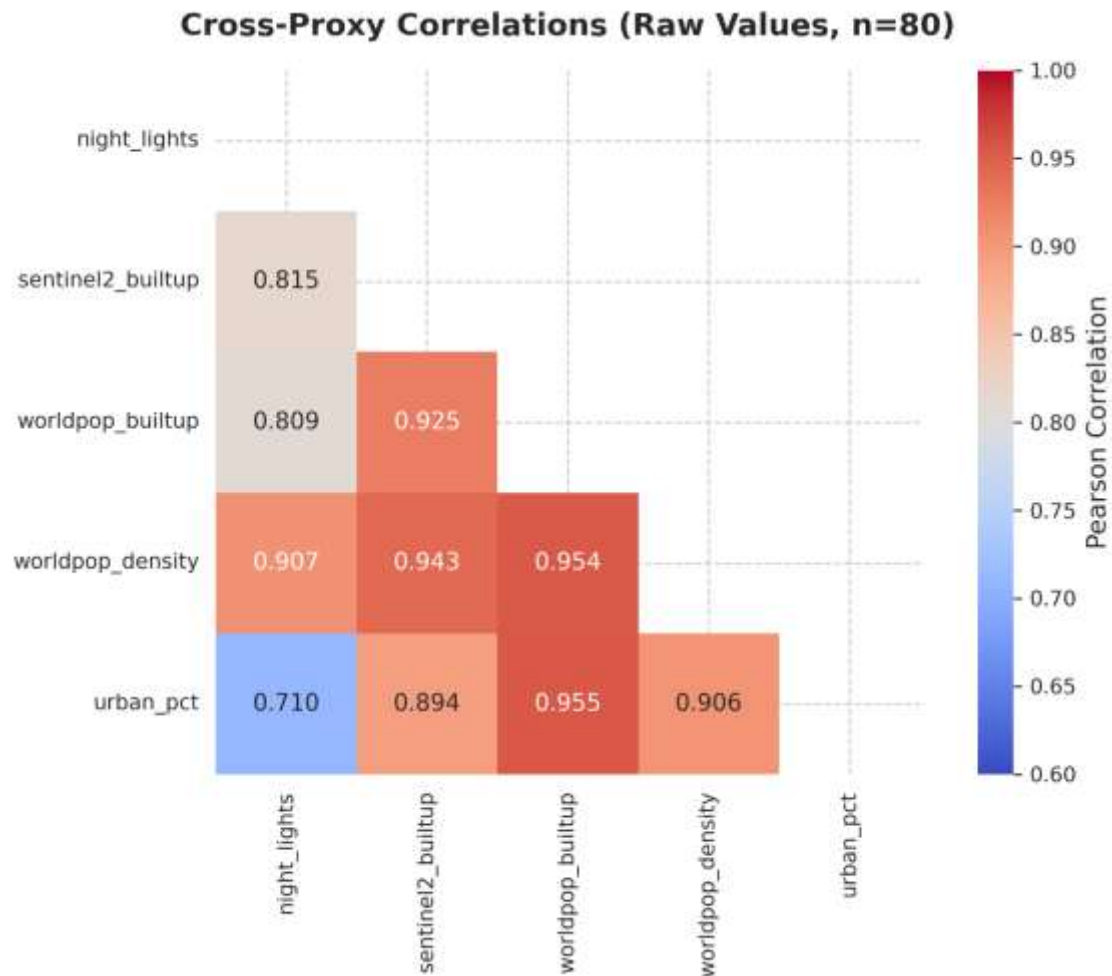


Figure 3. *Pearson correlation heatmap for pooled raw proxy values across the CHU-year panel.*

Notes: Correlations are calculated across 80 observations (10 CHUs across 8 years). Cell labels show Pearson correlation coefficients. Full corresponding 95% confidence intervals and p values are reported in Supplementary Tables S4A to S4C.

Independent k-means clustering with two clusters per proxy generated cross-proxy vote scores ranging from 0 to 5 (Figure 4; Supplementary Table S5). Three CHUs achieved consensus peri-urban classification under the prespecified threshold of 3 or higher: Buni and Mwele Kisurutini, each with a score of 5, and Vishakani, with a score of 4. Kombeni and Kwale each received intermediate scores of

2, while Mnazimwenga, Mutsengo, Viragoni, Kibwabwani, and Tsangatsini each had a score of 0 across all proxies. Overall, 3 of 10 CHUs (30.0%) were classified as peri-urban, with an exact binomial 95% confidence interval of 6.7% to 65.2%. Proxy-specific cluster assignments and vote score components are reported in Supplementary Table S2, and proxy-specific cluster visualisations are provided in Supplementary Figures S6A to S6E.

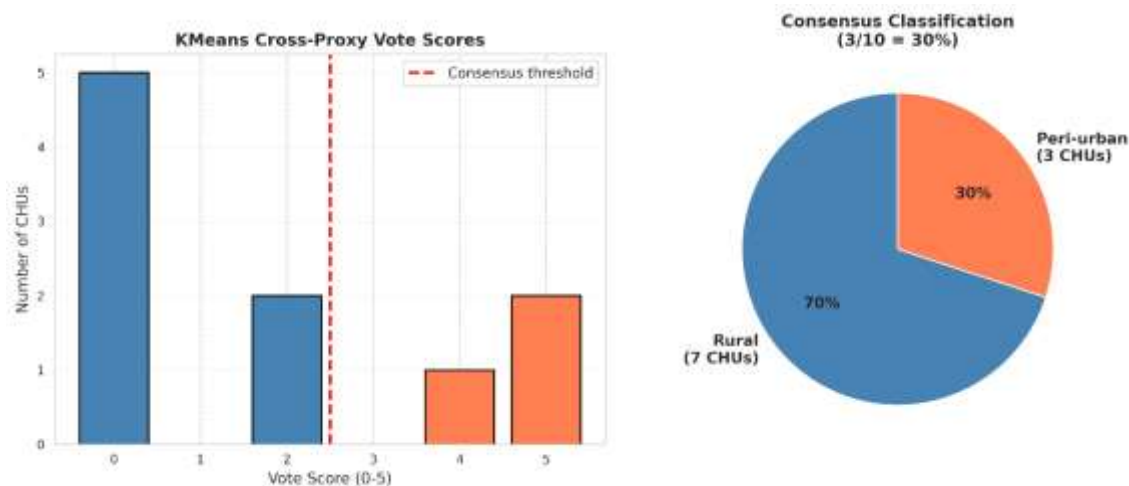


Figure 4. Cross-proxy vote scores and consensus peri-urban classification derived from *k*-means clustering. Notes: Vote score equals the number of proxies (from 0 to 5) for which a CHU is assigned to the higher intensity cluster. Consensus peri-urban is defined as vote score 3 to 5. The pie chart summarizes the proportion of CHUs classified as peri-urban under the prespecified rule.

Discussion

This study shows that a reproducible multi-proxy approach can be used to characterize settlement dynamics at CHU level within the KRHDSS. Three principal findings emerged. First, all five complementary proxies increased at system level from 2017 to 2024, but with divergent magnitudes, spatial patterns, and temporal stability. Second, the proxies were not interchangeable for temporal interpretation. Sentinel-2 built-up area showed substantial interannual volatility, whereas Degree of Urbanization showed a strong floor effect that reduced the number of defined trend observations. Third, cross-proxy consensus classification identified three CHUs, namely Buni, Mwele Kisurutini, and Vishakani, as consistently higher intensity across multiple dimensions. These findings reinforce the view that settlement transition is not a unitary process but a composite one in which physical expansion,

population concentration, and socioeconomic activity can evolve asynchronously in peri-urban settings [2-5,11,19].

The strong correlations between pooled raw proxy values suggest that the five indicators captured a common underlying gradient of settlement intensity. At the same time, the weak agreement between year-to-year and cumulative change metrics indicates that short-term temporal signals vary meaningfully by product design and data source. This distinction is important for health planning. A CHU may be consistently more urbanized overall even when individual proxies disagree about the pace or timing of recent change. Similar concerns have been raised in work examining the effects of gridded population model choice, the cell-level accuracy of population surfaces, and the implications of such differences for access and service analyses in sub-Saharan Africa [14,25-28].

The constrained growth range observed for WorldPop population density is plausible given the properties of gridded demographic modelling. Modelled population surfaces redistribute counts using covariates and smoothing procedures that can attenuate local extremes near settlement edges, especially in fragmented peri-urban landscapes [14,25-28]. By contrast, the volatility observed in Sentinel-2 built-up area likely reflects the limitations of optical classification in tropical settings, including cloud contamination, seasonal vegetation dynamics, and sensitivity to fragmented development patterns [29,30]. Nighttime lights occupied an intermediate position. It showed robust growth with relatively tight dispersion while identifying a smaller core of high-intensity CHUs than the built environment proxies. This is consistent with radiance functioning more as a proxy for electrification and economic activity than for physical structures alone [18,31-34]. Degree of Urbanization behaved differently again. Its structural zeros were informative for classification, but they constrained trend analysis and reduced sensitivity to early-stage transitions that remained below threshold cut-offs [19,35,36].

The consensus peri-urban classification is operationally useful as an exploratory framework. The three CHUs identified as peri-urban by cross-proxy vote scoring corresponded to the three units already recognized as peri-urban in operational work within the KRHDSS platform [37]. This concordance

provides face validity, but it should not be interpreted as formal external validation [9]. For community health planning, periodic updating of CHU settlement context could help refine workload estimation, outreach strategies, supervision intensity, and interpretation of routine service coverage indicators [6-9]. The approach may also help local programmes move beyond static rural-urban labels toward classifications that are more consistent with observed spatial change.

Several limitations warrant acknowledgement. First, the analytic sample comprised only 10 CHUs within one surveillance platform, limiting statistical precision and generalizability. Second, we analyzed aggregated CHU-level summaries rather than within-unit spatial heterogeneity. Third, some observed volatility almost certainly reflects source data properties rather than true settlement reversal. Fourth, the consensus classification depends on analyst-defined choices, including the use of $k = 2$, the median summary across 2017 to 2024, and the majority-vote threshold of 3 to 5, and alternative specifications could yield different CHU assignments. Fifth, built-up extent proxies capture the spatial extent of built surfaces rather than construction materials, housing quality, or service readiness. Sixth, although the consensus classification is internally coherent and operationally plausible, it was not validated against an independent external standard such as systematically assessed building footprints or field verification [38-45].

Future work should prioritize four directions. First, methodological refinements should incorporate stronger cloud masking and temporal smoothing for optical products in tropical settings [29,30,46,47]. Second, sensitivity analyses should test alternative clustering specifications and alternative vote thresholds. Third, validation should focus on feasible CHU-scale strategies, including periodic field verification and linkage to high-resolution building footprint data where available [38-42]. Fourth, routine repetition of this analysis within the KRHDSS could support earlier identification of emerging transitional CHUs and create a clearer evidence base for adaptive community health planning.

Conclusions

A multi-proxy geospatial approach can provide a reproducible and potentially useful way to characterize settlement context at CHU level in transitional settings. In this study, the approach revealed broad agreement across five proxies, but also important differences in their temporal behaviour. Combining these signals through cross-proxy vote scoring yielded an exploratory consensus classification that aligned with local operational understanding. Periodic updating of CHU settlement classification using open geospatial data could strengthen planning, targeting, and interpretation of routine community health indicators in rapidly changing environments, but further validation is needed before routine adoption at scale.

Abbreviations

CHU, Community Health Unit

DEGURBA, Degree of Urbanization

IQR, interquartile range

KRHDSS, Kaloleni Rabai Health and Demographic Surveillance System

QGIS, Quantum Geographic Information System

VIIRS, Visible Infrared Imaging Radiometer Suite

Declarations

Consent for publication

Not applicable.

Availability of data and materials

All input datasets used in this study are publicly available from the sources cited in the manuscript. The derived CHU-year analytical dataset and the full reproducible analysis code are deposited on Zenodo (DOI: <https://doi.org/10.5281/zenodo.19200614>) and are available from the corresponding authors on

reasonable request, subject to KRHDSS data-sharing requirements. Spatial boundary files for CHUs are available from the KRHDSS data management team subject to data sharing approval and use agreements and can be requested through the relevant data access process.

Competing interests

The authors declare no competing interests.

Funding

The KRHDSS platform receives core institutional support from Aga Khan University, supplemented by contributions from studies that utilize KRHDSS data and infrastructure. Aga Khan University had no role in the study design, data collection, data analysis, data interpretation, or writing of the manuscript.

Authors' contributions

PM and FO conceived the study, developed the methodology, curated and managed the data, performed the formal analysis, and drafted the original manuscript. RIO, FG, and AK contributed to the study design, assisted with interpretation of the results, and provided critical revisions of the manuscript. JD, ZB, and AN contributed to conceptualization of the study, offered methodological and policy guidance, and critically revised the manuscript for important intellectual content. All authors read and approved of the final manuscript.

Acknowledgements

We acknowledge the KRHDSS field and data management teams and the county and sub-county health leadership in Kilifi for their ongoing collaboration and operational insights.

References

1. United Nations, Department of Economic and Social Affairs, Population Division. World Urbanization Prospects 2018: Highlights. New York (NY): United Nations; 2019.
2. Combes PP, Gorin C, Nakamura S, Roberts M. An anatomy of urbanization in sub-Saharan Africa. Reg Sci Urban Econ. 2025;115:104152. doi:10.1016/j.regsciurbeco.2025.104152.

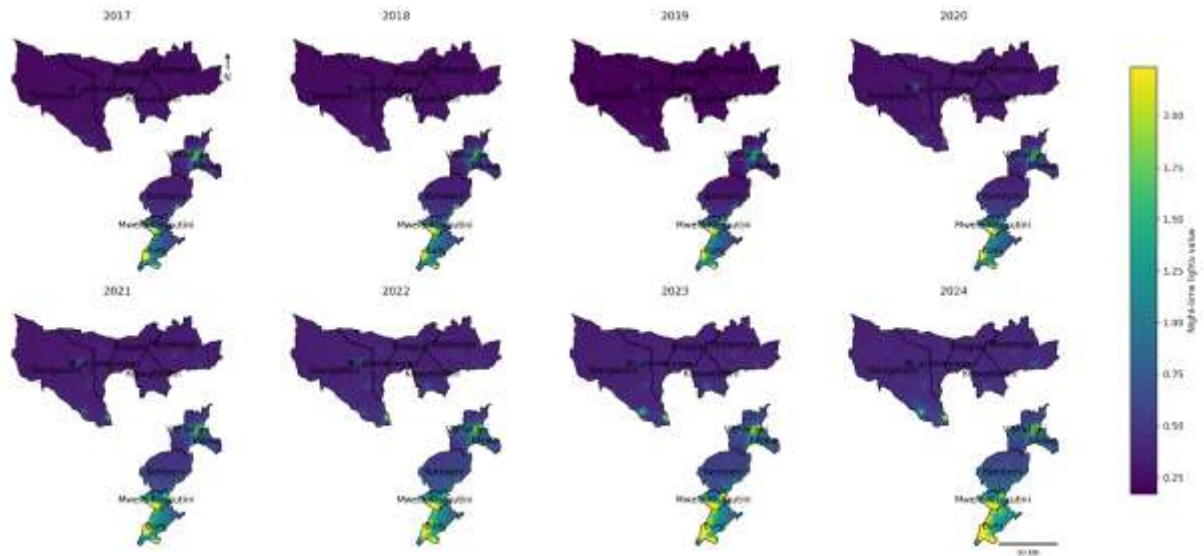
- 382 3. Saghir J, Santoro J. Urbanization in Sub-Saharan Africa: Meeting Challenges by Bridging
383 Stakeholders. Washington (DC): Center for Strategic and International Studies; 2018.
- 384 4. Agergaard J, Kirkegaard S, Birch-Thomsen T. Between village and town: small town urbanism in
385 sub-Saharan Africa. *Sustainability*. 2021;13(3):1417. doi:10.3390/su13031417.
- 386 5. Karg H, Hologa R, Schlesinger J, Drescher A, Kranjac-Berisavljevic G, Glaser R. Classifying and
387 mapping periurban areas of rapidly growing medium-sized sub-Saharan African cities: a multi-
388 method approach applied to Tamale, Ghana. *Land*. 2019;8(3):40. doi:10.3390/land8030040.
- 389 6. Wahl B, Lehtimäki S, Germann S, Schwalbe N. Expanding the use of community health workers in
390 urban settings: a potential strategy for progress towards universal health coverage. *Health Policy*
391 *Plan*. 2020;35(1):91–101. doi:10.1093/heapol/czz134.
- 392 7. Wafula CO, Edwards N, Kasibe DCO. Contextual variations in costs for a community health
393 strategy implemented in rural, peri-urban and nomadic sites in Kenya. *BMC Public Health*.
394 2017;17:224. doi:10.1186/s12889-017-4140-z.
- 395 8. Ministry of Health (Kenya). Kenya Community Health Strategy 2020–2025. Nairobi: Government
396 of Kenya; 2020.
- 397 9. Ngugi AK, Odhiambo R, Agoi F, Lakhani A, Orwa J, Obure J, et al. Cohort profile: the
398 Kaloleni/Rabai Community Health and Demographic Surveillance System. *Int J Epidemiol*.
399 2020;49(3):758–759e. doi:10.1093/ije/dyz252.
- 400 10. Hutchings P, Willcock S, Lynch K, et al. Understanding rural-urban transitions in the Global
401 South through peri-urban turbulence. *Nat Sustain*. 2022;5:924–930. doi:10.1038/s41893-022-
402 00920-w.
- 403 11. Samper J, Shelby JA, Behary D. The paradox of informal settlements revealed in an ATLAS of
404 informality: findings from mapping growth in the most common yet unmapped forms of
405 urbanization. *Sustainability*. 2020;12(22):9510. doi:10.3390/su12229510.
- 406 12. Pesaresi M, Schiavina M, Politis P, Freire S, Krasnodębska K, Uhl JH, et al. Advances on the
407 Global Human Settlement Layer by joint assessment of Earth observation and population survey
408 data. *Int J Digit Earth*. 2024;17(1):2390454. doi:10.1080/17538947.2024.2390454.
- 409 13. Palacios-Lopez D, Bachofer F, Esch T. High-resolution gridded population datasets: exploring the
410 capabilities of the World Settlement Footprint 2019 imperviousness layer for the African
411 continent. *Remote Sens*. 2021;13(6):1142. doi:10.3390/rs13061142.
- 412 14. Bondarenko M, Priyatikanto R, Tejedor Garavito N, Zhang W, McKeen T, Cunningham A, et al.
413 Constrained estimates of 2015–2030 total number of people per grid square at a resolution of 3
414 arc seconds, R2025A version v1 [dataset]. Southampton: University of Southampton; 2025.
415 doi:10.5258/SOTON/WP00839.
- 416 15. Lloyd CT, Sorichetta A, Tatem AJ. High resolution global gridded data for use in population
417 studies. *Sci Data*. 2017;4:170001. doi:10.1038/sdata.2017.1.
- 418 16. Woods T, McKeen T, Cunningham A, Priyatikanto R, Sorichetta A, Tatem AJ, et al. WorldPop
419 high resolution, harmonised annual global geospatial covariates. Version 1.0 [dataset].
420 Southampton: University of Southampton; 2025. doi:10.5258/SOTON/WP00772.
- 421 17. Karra K, Kontgis C, Statman-Weil Z, Mazzariello JC, Mathis M, Brumby SP. Global land
422 use/land cover with Sentinel-2 and deep learning. In: 2021 IEEE International Geoscience and
423 Remote Sensing Symposium; 2021 Jul 11–16; Brussels, Belgium. New York (NY): IEEE; 2021.
424 doi:10.1109/IGARSS47720.2021.9553499.

- 425 18. Elvidge CD, Zhizhin M, Ghosh T, Hsu FC, Taneja J. Annual time series of global VIIRS
426 nighttime lights derived from monthly averages: 2012 to 2019. *Remote Sens.* 2021;13(5):922.
427 doi:10.3390/rs13050922.
- 428 19. Bondarenko M, Zhang W, Tejedor Garavito N, Nosatiuk B, Tatem A, Sorichetta A. Degree of
429 Urbanization classes using WorldPop Global2 R2025A population data with the GHSL GHS-
430 DUG tool [dataset]. Southampton: University of Southampton; 2025.
431 doi:10.5258/SOTON/WP00879.
- 432 20. QGIS Development Team. qgis/QGIS: 3.28.3 [software]. Zenodo; 2023.
433 doi:10.5281/zenodo.7575913.
- 434 21. Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, et al. Scikit-learn:
435 machine learning in Python. *J Mach Learn Res.* 2011;12:2825–30.
- 436 22. Clopper CJ, Pearson ES. The use of confidence or fiducial limits illustrated in the case of the
437 binomial. *Biometrika.* 1934;26(4):404–13.
- 438 23. Efron B, Tibshirani RJ. *An Introduction to the Bootstrap.* Boca Raton (FL): Chapman &
439 Hall/CRC; 1994.
- 440 24. Davison AC, Hinkley DV. *Bootstrap Methods and Their Application.* Cambridge: Cambridge
441 University Press; 1997.
- 442 25. Hierink F, Boo G, Macharia PM, Ouma PO, Timoner P, Levy M, et al. Differences between
443 gridded population data impact measures of geographic access to healthcare in sub-Saharan
444 Africa. *Commun Med (Lond).* 2022;2:141. doi:10.1038/s43856-022-00179-4.
- 445 26. Thomson DR, Leasure DR, Bird T, Tzavidis N, Tatem AJ. How accurate are WorldPop-Global-
446 Unconstrained gridded population data at the cell-level? A simulation analysis in urban Namibia.
447 *PLoS One.* 2022;17(7):e0271504. doi:10.1371/journal.pone.0271504.
- 448 27. Zhang W, Woods D, Olowe ID, Schiavina M, Fang W, Hornby G, et al. Assessing the impacts of
449 gridded population model choice on degree of urbanization metrics. *Cities.* 2025;166:106293.
450 doi:10.1016/j.cities.2025.106293.
- 451 28. Thomson DR, Gaughan AE, Stevens FR, Yetman G, Elias P, Chen R. Evaluating the accuracy of
452 gridded population estimates in slums: a case study in Nigeria and Kenya. *Urban Sci.*
453 2021;5(2):48. doi:10.3390/urbansci5020048.
- 454 29. Tarrío K, Tang X, Masek JG, Claverie M, Ju J, Qiu S, et al. Comparison of cloud detection
455 algorithms for Sentinel-2 imagery. *Sci Remote Sens.* 2020;2:100009.
456 doi:10.1016/j.srs.2020.100009.
- 457 30. Huang C, Zhang C, He Y, Liu Q, Li H, Su F, et al. Land cover mapping in cloud-prone tropical
458 areas using Sentinel-2 data: integrating spectral features with NDVI temporal dynamics. *Remote*
459 *Sens.* 2020;12(7):1163. doi:10.3390/rs12071163.
- 460 31. Min B, Gaba KM, Sarr OF, Agalassou A. Detection of rural electrification in Africa using DMSP-
461 OLS night lights imagery. *Int J Remote Sens.* 2013;34(22):8118–41.
462 doi:10.1080/01431161.2013.833358.
- 463 32. Pérez-Sindín X, Chen T, Prishchepov A. Are night-time lights a good proxy of economic activity
464 in rural areas in middle and low-income countries? Examining the empirical evidence from
465 Colombia. *Remote Sens Appl Soc Environ.* 2021;24:100647. doi:10.1016/j.rsase.2021.100647.
- 466 33. Kabanda TH. Using land cover, population, and night light data to assess urban expansion in
467 Kimberley, South Africa. *S Afr Geogr J.* 2022;104(4):539–52.
468 doi:10.1080/03736245.2022.2028667.

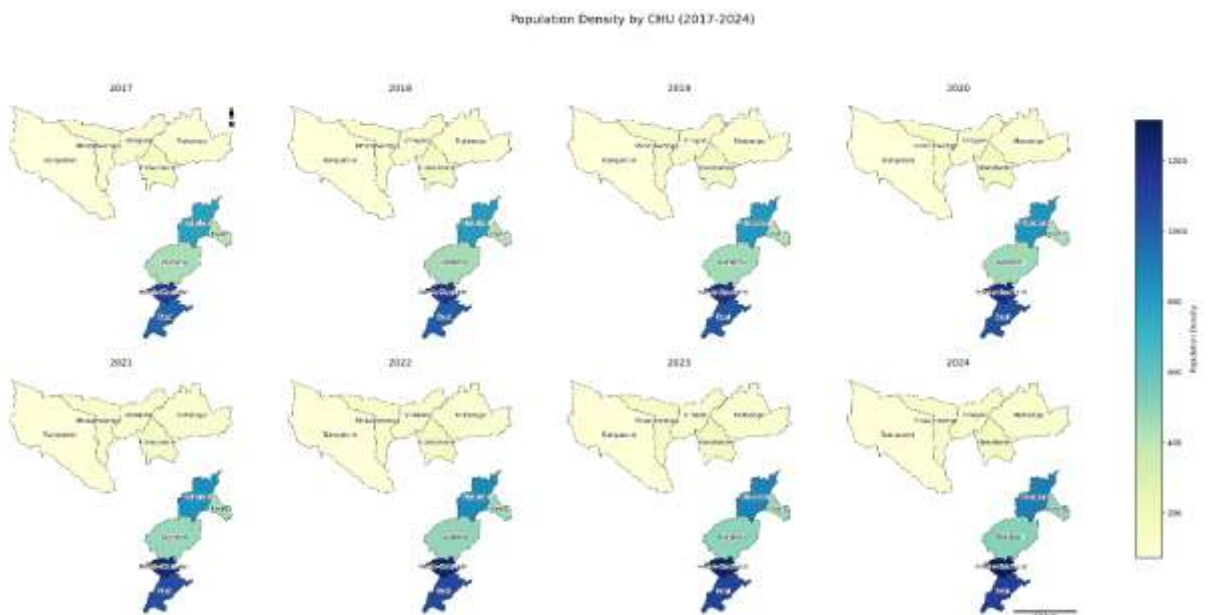
- 469 34. Singh S, Jain K, Shukla A. GIS integration of land cover with night-time lights for spatiotemporal
470 evaluation of urban expansion. *Int Arch Photogramm Remote Sens Spatial Inf Sci*. 2023;XLVIII-
471 1/W2-2023:1561–68. doi:10.5194/isprs-archives-XLVIII-1-W2-2023-1561-2023.
- 472 35. OECD, European Commission, Food and Agriculture Organization of the United Nations, United
473 Nations Human Settlements Programme, International Labour Organization, World Bank.
474 Applying the Degree of Urbanization: A Methodological Manual to Define Cities, Towns and
475 Rural Areas for International Comparisons. Paris: OECD Publishing; 2021.
476 doi:10.1787/4bc1c502-en.
- 477 36. Maffenini L, Schiavina M, Melchiorri M, Pesaresi M, Kemper T. GHS-DUG user guide: degree
478 of urbanization grid user guide. Version 4. Luxembourg: Publications Office of the European
479 Union; 2020. doi:10.2760/916217.
- 480 37. Ogero M, Orwa J, Odhiambo R, Agoi F, Lusambili A, Obure J, et al. Pentavalent vaccination in
481 Kenya: coverage and geographical accessibility to health facilities using data from a community
482 demographic and health surveillance system in Kilifi County. *BMC Public Health*.
483 2022;22(1):826. doi:10.1186/s12889-022-12570-w.
- 484 38. Oostwegel LJN, Schorlemmer D, Guéguen P, et al. From footprints to functions: a comprehensive
485 global and semantic building footprint dataset. *Sci Data*. 2025;12:1699. doi:10.1038/s41597-025-
486 06132-z.
- 487 39. Chamberlain HR, Darin E, Adewole WA, Jochem WC, Lázár AN, Tatem AJ. Building footprint
488 data for countries in Africa: to what extent are existing data products comparable? *Comput
489 Environ Urban Syst*. 2024;110:102104. doi:10.1016/j.compenvurbsys.2024.102104.
- 490 40. Hecht R, Kunze C, Hahmann S. Measuring completeness of building footprints in OpenStreetMap
491 over space and time. *ISPRS Int J Geo-Inf*. 2013;2(4):1066–91. doi:10.3390/ijgi2041066.
- 492 41. Ullah T, Lautenbach S, Herfort B, Reinmuth M, Schorlemmer D. Assessing completeness of
493 OpenStreetMap building footprints using MapSwipe. *ISPRS Int J Geo-Inf*. 2023;12(4):143.
494 doi:10.3390/ijgi12040143.
- 495 42. Sirko W, Kashubin S, Ritter M, Annkah A, Salah Y, Bouchareb E, et al. Continental-scale
496 building detection from high resolution satellite imagery. *arXiv [Preprint]*.
497 2021:arXiv:2107.12283.
- 498 43. Molla YB, Nilsen K, Singh K, Ruktanonchai CW, Schmitz MM, Duong J, et al. Best practices in
499 availability, management and use of geospatial data to guide reproductive, maternal, child and
500 adolescent health programmes. *BMJ Glob Health*. 2019;4(Suppl 5):e001406. doi:10.1136/bmjgh-
501 2019-001406.
- 502 44. Randriamihaja M, Ihantamalala FA, Rafenoarimalala FH, Finnegan KE, Rakotonirina L,
503 Razafinjato B, et al. Combining OpenStreetMap mapping and route optimization algorithms to
504 inform the delivery of community health interventions at the last mile. *PLOS Digit Health*.
505 2024;3(11):e0000621. doi:10.1371/journal.pdig.0000621.
- 506 45. Tatem AJ, Noor AM, Hay SI. Defining approaches to settlement mapping for public health
507 management in Kenya using medium spatial resolution satellite imagery. *Remote Sens Environ*.
508 2004;93(1–2):42–52. doi:10.1016/j.rse.2004.06.014.
- 509 46. Zhu Z, Wang S, Woodcock CE. Improvement and expansion of the Fmask algorithm: cloud,
510 cloud shadow, and snow detection for Landsats 4–7, 8, and Sentinel-2 images. *Remote Sens
511 Environ*. 2015;159:269–77. doi:10.1016/j.rse.2014.12.014.
- 512 47. Sanchez A, Picoli MCA, Camara G. Comparison of cloud cover detection algorithms on Sentinel-
513 2 images of the Amazon tropical forest. *Remote Sens*. 2020;12(8):1284. doi:10.3390/rs12081284.

Supplementary Material

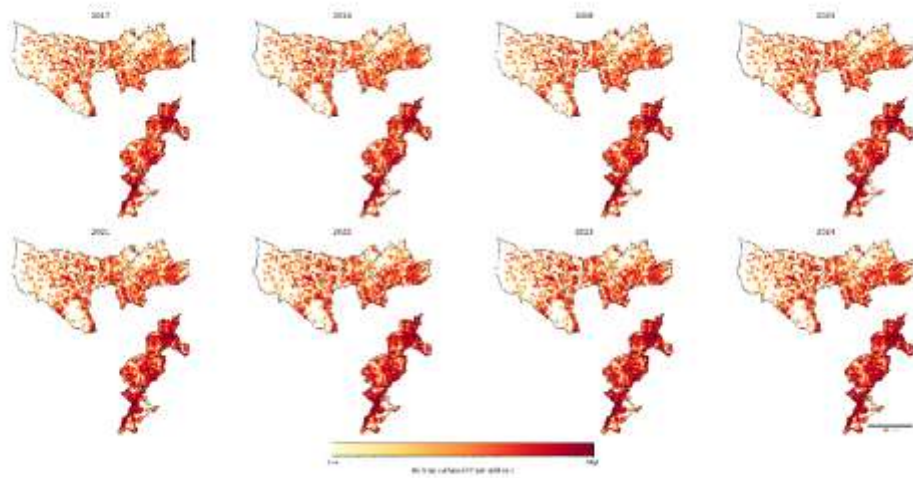
Supplementary Figures



Supplementary Figure S1. Annual spatial panels of nighttime lights intensity by CHU, 2017 to 2024. Notes: Annual values are derived from VIIRS nighttime lights and summarized to the CHU scale.

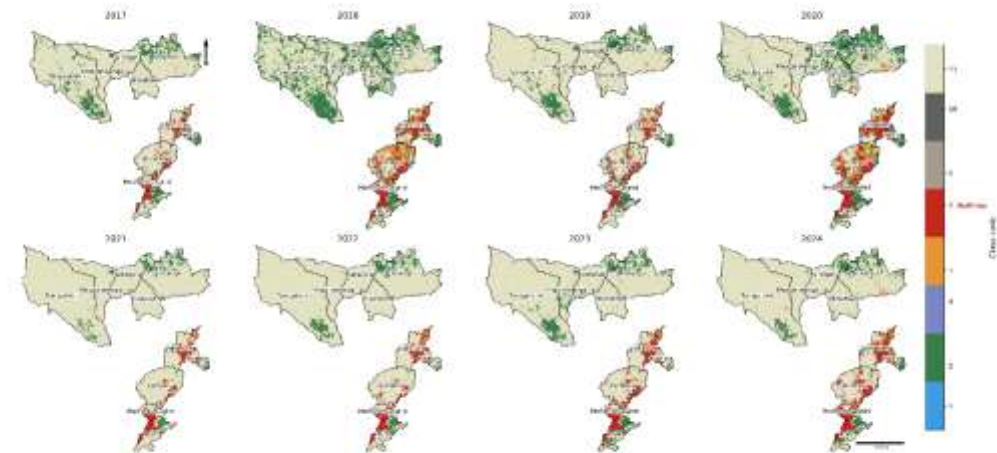


Supplementary Figure S2. Annual spatial panels of WorldPop population density by CHU, 2017 to 2024. Notes: Population density is expressed as persons per km² and summarized to the CHU scale.



Supplementary Figure S3. *Annual spatial panels of WorldPop built-up area by CHU, 2017 to 2024.*

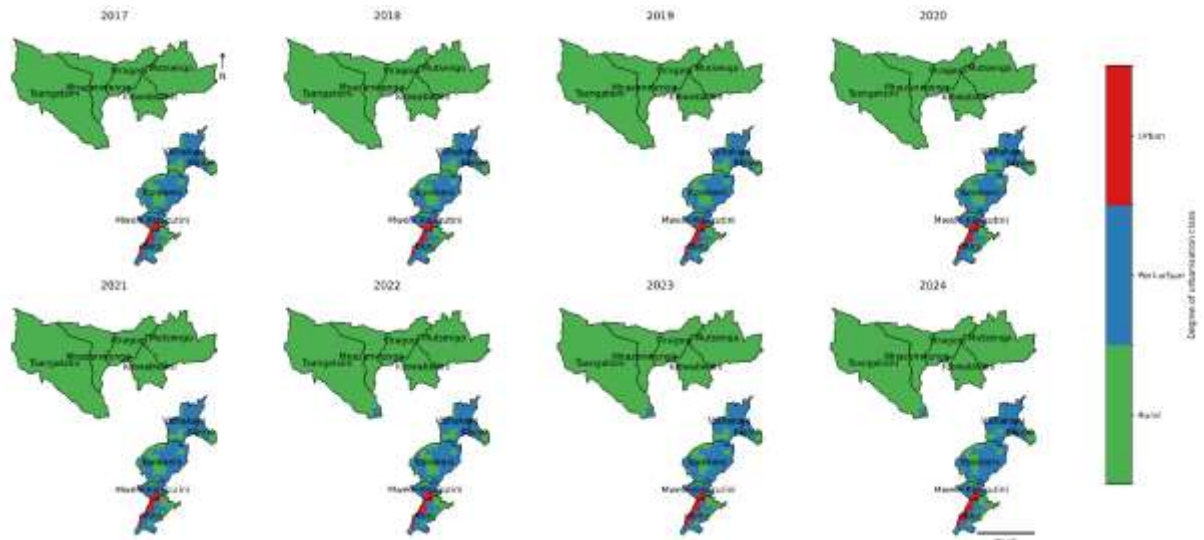
Notes: Built-up area is expressed as a percentage and summarized to the CHU scale.



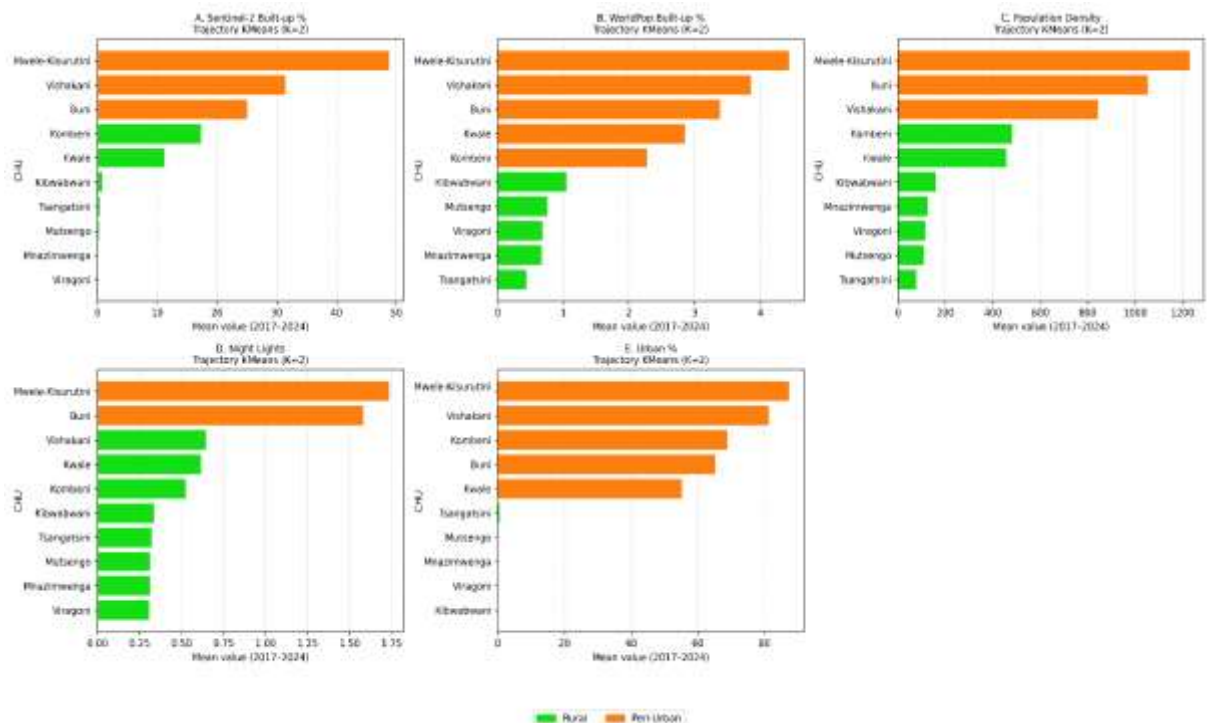
Supplementary Figure S4. *Annual spatial panels of Sentinel-2 built-up percentage by CHU, 2017 to*

2024. Notes: Apparent volatility should be interpreted in the context of small baseline values in some

CHUs, which can amplify percent change metrics.



Supplementary Figure S5. Annual spatial panels of Degree of Urbanization (urban percentage) by CHU, 2017 to 2024. Notes: The figure visualizes the floor effect, including CHUs with 0% urban percentage across all years.



Supplementary Figure S6 (A to E). Proxy-specific k-means classification using CHU mean values, 2017 to 2024. Panel A: Sentinel-2 built-up percentage; Panel B: WorldPop built-up percentage; Panel C: WorldPop population density; Panel D: nighttime lights intensity; Panel E: Degree of Urbanization

(urban percentage). Bar colour indicates cluster membership (lower-intensity: rural; higher-intensity: peri-urban). These panels provide the proxy-specific inputs used to construct the cross-proxy vote score.

Supplementary Tables

Supplementary Table S1a. CHU-level baseline, endline, absolute change, and cumulative percent change for nighttime lights intensity, 2017 to 2024.

CHU	2017 baseline	2024 endline	Absolute change	Cumulative % change
Buni	1.243	2.184	0.942	75.8%
Kibwabwani	0.286	0.437	0.152	53.2%
Kombeni	0.405	0.713	0.307	75.8%
Kwale	0.483	0.793	0.310	64.2%
Mnazimwenga	0.253	0.405	0.152	60.2%
Mutsengo	0.252	0.416	0.164	65.0%
Mwele Kisurutini	1.321	1.922	0.602	45.6%
Tsangatsini	0.256	0.437	0.181	70.6%
Viragoni	0.261	0.395	0.134	51.3%
Vishakani	0.536	0.810	0.274	51.2%

Notes: Cumulative percent change is undefined if the 2017 baseline equals zero.

Supplementary Table S1b. CHU-level baseline, endline, absolute change, and cumulative percent change for Sentinel-2 built-up percentage, 2017 to 2024.

CHU	2017 baseline	2024 endline	Absolute change	Cumulative % change
Buni	20.998	28.417	7.419	35.3%
Kibwabwani	0.001	0.392	0.391	30175.0%
Kombeni	10.513	22.155	11.642	110.7%
Kwale	6.135	12.412	6.276	102.3%
Mnazimwenga	0.004	0.027	0.022	505.6%
Mutsengo	0.105	0.335	0.230	218.9%
Mwele Kisurutini	42.074	58.637	16.563	39.4%
Tsangatsini	0.189	0.426	0.237	125.1%
Viragoni	0.000	0.001	0.001	
Vishakani	25.896	32.454	6.558	25.3%

Notes: Large cumulative percent changes can occur when the 2017 baseline is close to zero. Cumulative percent change is undefined where the 2017 baseline equals zero.

Supplementary Table S1c. CHU-level baseline, endline, absolute change, and cumulative percent change for WorldPop built-up area, 2017 to 2024.

CHU	2017 baseline	2024 endline	Absolute change	Cumulative % change
Buni	3.07	3.677	0.607	19.8%
Kibwabwani	0.867	1.239	0.372	42.9%
Kombeni	1.777	2.766	0.989	55.7%
Kwale	2.365	3.386	1.021	43.2%
Mnazimwenga	0.589	0.761	0.172	29.2%
Mutsengo	0.583	0.930	0.347	59.5%
Mwele Kisurutini	3.874	4.952	1.078	27.8%
Tsangatsini	0.397	0.486	0.089	22.4%
Viragoni	0.581	0.778	0.197	33.9%
Vishakani	3.291	4.422	1.131	34.4%

Notes: Built-up area expressed as percentage. All CHUs have non-zero 2017 baselines.

Supplementary Table S1d. CHU-level baseline, endline, absolute change, and cumulative percent change for WorldPop population density, 2017 to 2024.

CHU	2017 baseline	2024 endline	Absolute change	Cumulative % change
Buni	985.1	1123.8	138.7	14.1%
Kibwabwani	147.8	168.9	21.1	14.3%
Kombeni	446.4	515.9	69.5	15.6%
Kwale	421.4	495.3	73.9	17.5%
Mnazimwenga	118.9	134.1	15.2	12.8%
Mutsengo	102.0	117.7	15.8	15.5%
Mwele Kisurutini	1143.1	1315.2	172.1	15.1%
Tsangatsini	73.4	83.2	9.8	13.4%
Viragoni	107.9	122.0	14.1	13.1%
Vishakani	785.3	904.9	119.7	15.2%

Notes: Population density expressed as persons per km².

Supplementary Table S1e. CHU-level baseline, endline, absolute change, and cumulative percent change for Degree of Urbanization (urban percentage), 2017 to 2024.

CHU	2017 baseline	2024 endline	Absolute change	Cumulative % change
Buni	65.02%	65.34%	0.31 pp	0.5%
Kibwabwani	0.00%	0.00%	0.00 pp	
Kombeni	65.44%	74.98%	9.54 pp	14.6%
Kwale	52.69%	59.62%	6.93 pp	13.2%

CHU	2017 baseline	2024 endline	Absolute change	Cumulative % change
Mnazimwenga	0.00%	0.00%	0.00 pp	
Mutsengo	0.00%	0.00%	0.00 pp	
Mwele Kisurutini	87.36%	87.36%	0.00 pp	0.0%
Tsangatsini	0.00%	1.18%	1.18 pp	
Viragoni	0.00%	0.00%	0.00 pp	
Vishakani	77.01%	89.06%	12.05 pp	15.7%

Notes: Absolute change in percentage points (pp). Cumulative percent change undefined for CHUs with 0% baseline.

Supplementary Table S2. Proxy-specific cluster assignments and vote score components used to derive cross-proxy consensus classification.

CHU	NTL	Sentinel-2	WP built up	WP density	Urban %	Score	Classification
Buni	1	1	1	1	1	5	Peri-urban
Kibwabwani	0	0	0	0	0	0	Rural
Kombeni	0	0	1	0	1	2	Rural
Kwale	0	0	1	0	1	2	Rural
Mnazimwenga	0	0	0	0	0	0	Rural
Mutsengo	0	0	0	0	0	0	Rural
Mwele Kisurutini	1	1	1	1	1	5	Peri-urban
Tsangatsini	0	0	0	0	0	0	Rural
Viragoni	0	0	0	0	0	0	Rural
Vishakani	0	1	1	1	1	4	Peri-urban

Notes: Higher-intensity cluster = 1; lower-intensity = 0. NTL = nighttime lights; WP = WorldPop. Vote score is the sum across proxies.

Supplementary Table S3. Availability of cumulative percent change outcomes from 2017 to 2024 by proxy.

Proxy	Total possible CHUs	Defined CHUs	Undefined CHUs	Reason undefined
Nighttime lights intensity	10	10	0	
Sentinel-2 built up	10	9	1	Baseline equals 0 for 1 CHU
WorldPop built-up area	10	10	0	
WorldPop population density	10	10	0	
Urban percentage	10	5	5	Baseline equals 0 for 5 CHUs

Notes: Cumulative percent change undefined when the 2017 baseline equals zero.

Supplementary Table S4A. Pearson correlations between proxies using pooled raw values (80 CHU-year observations).

Proxy pair	r (95% CI)	p-value	N
Nighttime lights vs Sentinel-2 built up	0.815 (0.725, 0.877)	<0.0001	80
Nighttime lights vs WorldPop built-up area	0.809 (0.717, 0.874)	<0.0001	80
Nighttime lights vs WorldPop population density	0.907 (0.858, 0.939)	<0.0001	80
Nighttime lights vs urban percentage	0.710 (0.581, 0.804)	<0.0001	80
Sentinel-2 built up vs WorldPop built-up area	0.925 (0.885, 0.951)	<0.0001	80
Sentinel-2 built up vs WorldPop population density	0.943 (0.912, 0.963)	<0.0001	80
Sentinel-2 built up vs urban percentage	0.894 (0.840, 0.931)	<0.0001	80
WorldPop built-up area vs WorldPop population density	0.954 (0.929, 0.970)	<0.0001	80
WorldPop built-up area vs urban percentage	0.955 (0.931, 0.971)	<0.0001	80
WorldPop population density vs urban percentage	0.906 (0.856, 0.939)	<0.0001	80

Notes: All p values below 0.0001.

Supplementary Table S4B. Pearson correlations between proxies using pooled year-to-year percent changes across defined transitions.

Proxy pair	r (95% CI)	p-value	N
Nighttime lights vs Sentinel-2 built up	−0.095 (−0.339, 0.161)	0.4668	61
Nighttime lights vs WorldPop built-up area	−0.019 (−0.253, 0.217)	0.8764	70
Nighttime lights vs WorldPop population density	−0.046 (−0.278, 0.191)	0.7057	70
Nighttime lights vs urban percentage	0.189 (−0.144, 0.483)	0.2638	37
Sentinel-2 built up vs WorldPop built-up area	0.076 (−0.179, 0.322)	0.5586	61
Sentinel-2 built up vs WorldPop population density	0.104 (−0.151, 0.347)	0.4230	61
Sentinel-2 built up vs urban percentage	−0.039 (−0.359, 0.288)	0.8165	37
WorldPop built-up area vs WorldPop population density	0.484 (0.281, 0.646)	<0.0001	70
WorldPop built-up area vs urban percentage	0.186 (−0.147, 0.481)	0.2709	37
WorldPop population density vs urban percentage	0.243 (−0.088, 0.526)	0.1475	37

Notes: Only WorldPop built-up area and population density show a statistically significant correlation ($p < 0.0001$).

580

581 **Supplementary Table S4C.** Pearson correlations between proxies using CHU-level cumulative
 582 percent changes.

Proxy pair	r (95% CI)	p-value	N
Nighttime lights vs Sentinel-2 built up	−0.319 (−0.811, 0.438)	0.4028	9
Nighttime lights vs WorldPop built-up area	0.116 (−0.554, 0.695)	0.7501	10
Nighttime lights vs WorldPop population density	0.097 (−0.567, 0.685)	0.7896	10
Nighttime lights vs urban percentage	0.103 (−0.857, 0.903)	0.8689	5
Sentinel-2 built up vs WorldPop built-up area	0.153 (−0.569, 0.742)	0.6952	9
Sentinel-2 built up vs WorldPop population density	−0.151 (−0.741, 0.570)	0.6974	9
Sentinel-2 built up vs urban percentage	0.492 (−0.689, 0.958)	0.3993	5
WorldPop built-up area vs WorldPop population density	0.552 (−0.119, 0.877)	0.0979	10
WorldPop built-up area vs urban percentage	0.782 (−0.323, 0.985)	0.1180	5
WorldPop population density vs urban percentage	0.573 (−0.626, 0.967)	0.3128	5

583 Notes: Confidence intervals are wide owing to small effective sample sizes. No correlations are statistically significant at
 584 $p < 0.05$.

585

586 **Supplementary Table S5.** Cross-proxy vote scores and consensus classification from k-means
 587 clustering applied independently to each proxy.

CHU	Vote score (0 to 5)	Consensus classification
Buni	5	Peri-urban
Mwele Kisurutini	5	Peri-urban
Vishakani	4	Peri-urban
Kombeni	2	Rural
Kwale	2	Rural
Kibwabwani	0	Rural
Mutsengo	0	Rural
Mnazimwenga	0	Rural
Tsangatsini	0	Rural
Viragoni	0	Rural

588 Notes: Consensus peri-urban is defined as vote score 3 to 5.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [SupplementaryMaterialCHUUrbanization26032026FO.docx](#)