International Journal of Computer Applications Technology and Research
Volume 07–Issue 12, 449 - 472, 2018, ISSN:-2319–8656

Extraction and Validation of Database of Urban and Non-Urban Points from Remote Sensing Data

Olawale Luqman Ajani
Department: Electronic Engineering,
University of Pavia,
Italy

Abstract: Recent advancements in global settlement mapping have ushered in a new era of spatial analysis. The introduction of the Global Urban Footprint (GUF)—derived from radar-based TanDEM-X data—and the Global Human Settlement Layer (GHSL)—based on optical satellite data—offers unprecedented spatial resolution for mapping complex human settlement patterns. Despite these innovations, comparative assessments evaluating their performance against existing lower-resolution datasets remain limited. To address this, we present a robust cross-comparison framework that examines inter-map agreement and classification accuracy across multiple African landscapes with varying settlement typologies. Using a curated set of reference data points from thirteen African countries, the study assesses the concordance between ground observations and GUF outputs. Key indicators of accuracy, including the Kappa coefficient, overall agreement rates, and pattern-based metrics, are used to evaluate the performance of these new layers. Findings reveal that the high-resolution GUF and GHSL products offer significant improvements in delineating low-density and peri-urban settlements—areas previously underrepresented by coarser-resolution maps. Pattern-based analysis shows that accuracy is positively correlated with settlement structure, particularly in zones characterized by medium and small patch sizes. The results indicate a marked shift in the spatial fidelity of global settlement representations, particularly beyond urban core areas. This comprehensive validation supports the integration of GUF and GHSL data into urban planning and policy frameworks, enhancing the capacity to monitor and manage urbanization at multiple scales.

Keywords: Global Urban Footprint, Human Settlement Layer, Urban Mapping, Kappa Coefficient, Spatial Accuracy, Remote Sensing Africa

1. INTRODUCTION

1.1 Urban Areas and the Need to Use Satellite Data

The definition of "urban" is inherently relative, varying across countries and, at times, within the same country due to administrative reclassification. This variability complicates cross-national comparisons and temporal assessments of urbanization patterns [1]. Urban areas are generally characterized by high population density and significant concentrations of built infrastructure. They emerge through urbanization and manifest in various morphological forms such as cities, towns, suburbs, or conurbations.

A settlement may be classified as urban based on several criteria, including population size (often with thresholds ranging from 200 to 50,000 people), population density, economic structure (e.g., where the majority of residents are engaged in non-agricultural sectors), and the presence of infrastructural features such as paved roads, electric lighting, and sewerage systems [2]. Urban areas are typically heterogeneous, hosting populations diverse in race, ethnicity, religion, socioeconomic status, and occupational background. This diversity, coupled with high mobility among urban dwellers, creates intense competition for space and resources.

According to the World Bank, no region is urbanizing more rapidly than Africa. The continent's urban population, which stood at 36% in 2010, is projected to reach 50% by 2030 [3]. This rapid growth presents both opportunities and challenges. Urbanization can catalyze economic development, structural transformation, and poverty alleviation, but it also necessitates accurate spatial data for planning, monitoring, and managing urban expansion.

Given the challenges associated with inconsistent definitions and limited census data, satellite-based remote sensing has become essential in monitoring urban growth. Satellite imagery provides a consistent, objective, and scalable means of capturing urban dynamics, especially in regions where administrative data are sparse or outdated [4]. As African cities continue to expand at unprecedented rates, the integration of satellite data into urban planning processes is critical for achieving sustainable development outcomes [5].

Figure 1 Aerial view of the district of Plateau in Abidjan, Cote d'Ivoire (UN Photo/Basile Zoma)

1.2 Importance of Urban vs. Non-Urban Data in Spatial Planning

Differentiating between urban and non-urban areas is vital for spatial planning because it directly influences infrastructure investment, service provision, and environmental management. Urban data informs the design of transportation corridors, zoning regulations, and housing developments. Conversely, non-urban classifications guide agricultural planning, biodiversity conservation, and rural development initiatives (6).

In rapidly transforming geographies, especially in Sub-Saharan Africa, urban sprawl often overlaps with informal or unregulated land uses. Distinguishing these through updated satellite-derived data supports targeted interventions. The GUF and GHSL (Global Human Settlement Layer) datasets, for instance, offer high-resolution basemaps that reveal human settlement footprints even in low-density areas previously underrepresented in legacy mapping systems (7).

Such distinctions are not merely academic—they are necessary for determining eligibility for urban funding programs, assessing flood risk exposure in peri-urban zones, and enforcing land-use regulations. Urban-rural gradients also inform socio-economic programs by correlating settlement density with poverty indices, healthcare accessibility, and educational attainment levels (8).

Accurate urban-non-urban delineation thus serves as a foundation for resilient and equitable spatial development strategies (9).

1.3 Challenges in Accurate Extraction and Classification

Despite technical advancements, challenges in classification remain, particularly when distinguishing complex settlement forms. Urban areas are often characterized by heterogeneity—dense cores, scattered suburbs, and transitional zones coexisting within the same spatial tile. Traditional single-class pixel labelling often fails to capture this nuance, leading to misclassification and reduced reliability of derived statistics (10).

Additionally, issues arise from inconsistent imagery coverage, atmospheric interference, and mixed pixels in moderate-resolution datasets. For example, radar backscatter intensity used in GUF may overestimate built-up areas in regions with metallic roofs or underrepresent shaded or vegetated settlements (11). The absence of standard definitions of "urban" across nations further complicates classification comparability (12).

Validation remains a core bottleneck. Ground truthing is resource-intensive, especially in geographically expansive or politically unstable regions. The manual selection of training samples, as was required in Ajani's work, introduces subjectivity and scale-dependent bias (13). Differences in map legends, projection systems, and spatial resolutions across datasets also contribute to inter-map disagreement, limiting their interoperability.

Such challenges underscore the necessity for standardized, high-resolution, and regularly updated classification systems that balance computational efficiency with contextual sensitivity (14).

1.4 Research Aims and Contribution This study aims to extract and validate a database of urban and non-urban points from remotely sensed satellite data, focusing on multiple African countries as case examples. Through cross-comparison of classified urban extents—particularly using the Global Urban Footprint (GUF)—and visually validated reference points, the research evaluates classification accuracy and consistency (15).

Key contributions include the development of a scalable methodology for selecting statistically significant validation points across geographies, segmented into categories such as 25%, 50%, 75%, and 100% urbanized areas. The manual refinement of classification through expert consensus enhances credibility in cases of spectral ambiguity (16).

Furthermore, by quantifying inter-map agreement via confusion matrices and kappa coefficients, the research adds a valuable performance benchmark for future classification models. It also highlights the limitations of existing urban products in capturing low-density or peri-urban settlements—insights critical for refining future land monitoring frameworks (17).

2. REMOTE SENSING DATA FOR URBAN MAPPING

2.1 Overview of Remote Sensing Platforms

Remote sensing platforms have played a crucial role in the advancement of land classification, particularly for differentiating urban from non-urban regions. Three core platforms—**Landsat**, **Sentinel**, and **MODIS**—have been central to mapping land surface dynamics over time.

The **Landsat program**, initiated in the 1970s, marked a significant step forward by introducing the Multispectral Scanner (MSS) and later the Thematic Mapper (TM), which allowed for a 30-meter spatial resolution across visible, near-infrared, and shortwave infrared bands (7). With the Enhanced Thematic Mapper Plus (ETM+) and the Operational Land Imager (OLI), Landsat evolved into a widely used standard for urban growth analysis (8).

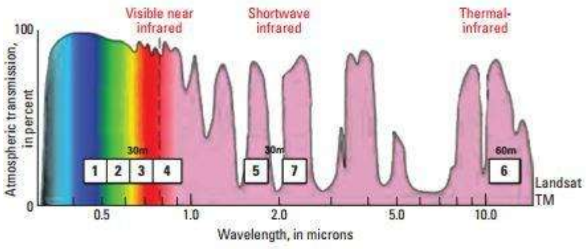

Figure 2: Locations of Landsat bands in the electromagnetic spectrum.

The **Sentinel missions**, especially Sentinel-1 and Sentinel-2 under the Copernicus programme, added robust capabilities to this domain. Sentinel-2 introduced 13 spectral bands and higher spatial resolution (up to 10 meters in some bands), which facilitated better vegetation, soil, and urban classification under varying atmospheric conditions (9). Sentinel-1, using C-band Synthetic Aperture Radar (SAR), extended mapping possibilities to cloudy or night-time conditions, especially in tropical regions where optical imagery suffers limitations (10).

MODIS (Moderate Resolution Imaging Spectroradiometer), with a coarser resolution (250m–1km) but daily revisit time, proved essential for detecting land cover change at continental and temporal scales (11). Although not optimal for fine-scale urban classification, MODIS is commonly used for trend monitoring and spectral index calculations due to its consistency and spectral richness.

Together, these platforms offer complementary advantages: **Landsat** for long-term and medium-resolution analysis, **Sentinel** for higher resolution and spectral granularity, and **MODIS** for high-frequency temporal observation (12).

2.2 Data Preprocessing Techniques

Preprocessing remote sensing data is essential to produce reliable, usable imagery for urban classification. This stage includes atmospheric correction, geometric rectification, cloud masking, and noise filtering, all of which mitigate distortions introduced during image acquisition.

Atmospheric correction adjusts for distortions caused by aerosols, gases, and water vapor, which affect the spectral reflectance received by sensors. One common method, Dark Object Subtraction (DOS), assumes certain land surfaces (e.g., deep water bodies) should ideally appear near zero reflectance in specific bands, thus helping estimate atmospheric haze (13).

Geometric rectification involves aligning satellite imagery with known ground control points to ensure positional accuracy. This is especially important when integrating images from different dates or platforms. Inaccurate co-registration can lead to misclassification and spatial misalignment in change detection studies (14).

Cloud masking is a critical step in tropical or humid regions. Algorithms like Fmask detect cloud and shadow pixels based on thermal and visible band thresholds and flag them for removal. Unmasked clouds can significantly mislead classification algorithms by mimicking high-reflectance surfaces such as concrete or metal roofs (15).

Noise filtering helps to smooth pixel-based artifacts introduced by sensor defects, scan line errors, or random reflectance variations. Techniques like low-pass convolution filters or kernel smoothing are applied to reduce spatial noise. In urban analysis, smoothing helps clarify boundaries between dense and low-density zones without compromising spatial detail (16).

In this project, specific preprocessing included bilinear downsampling of GUF's 12.5 m resolution to a 300 m scale, combined with 25x25 kernel smoothing to generalize classification areas. This combination helped maintain classification integrity while aligning data with coarser-resolution references like the GHSL (17).

Robust preprocessing ensures that classification inputs are clean, normalized, and reliable across geographies and timeframes.

2.3 Limitations of Raw Remote Sensing Data

While remote sensing data offers valuable insights, it is not without limitations—especially when used for fine-grained land classification across heterogeneous regions. These constraints arise from trade-offs in spatial resolution, sensor performance, and seasonal or atmospheric variability.

A key limitation is the resolution-versus-coverage trade-off. High-resolution images (e.g., <10 m) provide excellent urban detail but are typically limited in coverage, revisit time, and data volume. Conversely, coarser-resolution imagery (e.g., MODIS) enables broad-area and temporal studies but lacks the granularity to detect small-scale urban expansion or informal settlements (18).

Sensor-specific biases also impact accuracy. For example, SAR imagery from Sentinel-1 or TerraSAR-X may misinterpret bright returns from non-urban metallic or rocky features as built-up zones. Conversely, optical sensors like Landsat and Sentinel-2 are vulnerable to sun angle effects, sensor saturation, or obscuration by vegetation in peri-urban areas (19).

Another common issue is temporal inconsistency. Seasonal land cover changes (e.g., bare agricultural fields during dry seasons) may be misclassified as impervious surfaces. Urban classification models that do not account for vegetation cycles or land-use seasonality risk producing false positives or negatives (20).

Sensor malfunctions and anomalies—like the Landsat 7 scan line corrector (SLC) failure—can result in data gaps or striping, which must be compensated through image mosaicking or data fusion. Yet these repairs introduce further complexity to classification pipelines (21).

Finally, classification interoperability is challenged by non-standardized definitions of "urban" and differing land use legends across datasets. In this project, inter-map comparisons between GUF, GHSL, and GlobeLand30 revealed substantial discrepancies due to differing resolutions, urban criteria, and extraction methods (22). These inconsistencies make meta-analyses and cross-country comparisons difficult.

Therefore, while remote sensing is indispensable for urban-non-urban analysis, its full potential is only realized through proper interpretation, harmonization, and validation.

Table 1: Comparison of Satellite Platforms Used for Urban Classification

Satellite	**Sensor Type**	**Spatial Resolution**	**Spectral Bands**	**Temporal Revisit**	**Strengths**
Landsat 7/8	Optical (ETM+/OLI)	15–30 m	Visible, NIR, SWIR	16 days	Long archive; good for trend analysis
Sentinel-2	Optical (MSI)	10–60 m	13 VNIR/SWIR bands	5 days	High resolution; atmospheric clarity
Sentinel-1	Radar (C-SAR)	10–20 m	N/A (backscatter only)	6 days	Works in clouds and night
MODIS	Optical	250 m – 1 km	36 bands (varied widths)	1–2 days	Ideal for regional temporal analysis
GUF (DLR)	Radar (TanDEM-X)	12.5 m	Single radar frequency	Varies	High-resolution urban footprints

3. URBAN AND NON-URBAN CLASSIFICATION TECHNIQUES

3.1 Spectral Indices and Thresholding

Spectral indices have long served as efficient tools in land use and land cover (LULC) classification, especially in urban analysis. Among the most widely adopted indices is the Normalized Difference Vegetation Index (NDVI), which utilizes the difference between near-infrared and red reflectance to detect vegetation health and presence. Its simplicity and robustness in delineating vegetative areas made it a staple for satellite imagery interpretation during earlier remote sensing applications. NDVI thresholds were commonly defined based on empirical observation or expert input to separate vegetated from non-vegetated land surfaces, supporting coarse and fine-scale urban mapping tasks (11).

In urban studies, another pivotal index is the Normalized Difference Built-up Index (NDBI), which exploits the contrast between shortwave infrared and near-infrared bands. NDBI helps to emphasize built-up surfaces and, in combination with NDVI, enables binary classifications distinguishing urban fabric from vegetation or water. These binary classifications formed the foundational logic in pre-deep learning workflows for urban monitoring, facilitating structured LULC mapping workflows (12).

The Urban Index (UI), a combination of multiple spectral bands, aimed to further enhance urban surface extraction by minimizing the confusion between impervious surfaces and bare land. Its design often incorporated intermediate values from NDVI and NDBI to improve robustness in semi-urban or peri-urban landscapes, reducing errors common in single-index approaches (13). Thresholding methods, often informed by histograms or unsupervised clustering, allowed rapid segmentation, suitable for large datasets or environments with limited computational capacity.

While these indices provided computationally efficient methods of image classification, they inherently lacked context sensitivity, often confusing spectrally similar features such as bare soil and built-up land. To counter this, hybrid thresholding strategies were sometimes used, combining static thresholds with auxiliary spatial rules, particularly in urban zones with mixed land types (14). These approaches demonstrated particular value in low-resource settings where access to training data or high-end processing power was restricted.

Despite the eventual shift toward machine learning-based classification, spectral indices and thresholding continue to serve as a pre-processing or feature extraction step. They remain embedded in rule-based systems and as priors in more complex classification models, particularly those involving temporal or multisensor data integration (15).

3.2 Machine Learning and AI-Based Classification

Before the widespread use of deep learning, classical machine learning models were integral in the automation of LULC classification. Among these, the Support Vector Machine (SVM) was recognized for its high generalization capability and suitability for small training samples. SVMs used kernel functions to transform spectral features into higher-dimensional spaces, making them capable of resolving nonlinear class boundaries, especially in heterogeneous urban landscapes (16).

Random Forests (RF) emerged as another powerful classifier due to their ensemble nature, leveraging the outputs of multiple decision trees trained on random feature subsets. This approach minimized overfitting while maximizing classification accuracy across various land cover types. RFs also provided insight into feature importance, a trait that proved valuable when dealing with multispectral datasets and spectral indices. Their interpretability and robustness to noise contributed to their prominence in earlier geospatial analysis pipelines (17).

In parallel, advances in computer vision began to influence remote sensing through the application of Convolutional Neural Networks (CNNs). While computationally intensive, early CNN models offered superior performance in tasks requiring spatial pattern recognition. Their multi-layered architecture allowed hierarchical extraction of spatial features, which proved advantageous in mapping urban areas characterized by complex textures and geometries. The transition to CNNs was gradual, often initiated in small-scale experiments on urban patches (18).

A key benefit of these AI-based methods was their ability to generalize across sensors and scenes without extensive manual tuning. They replaced manual thresholding with data-driven decision boundaries, automatically learning patterns from labeled datasets. However, during this period, access to high-resolution training data and computational infrastructure was limited, which constrained the widespread use of CNNs outside of research settings (19).

Machine learning models also allowed the incorporation of multi-source data, such as LiDAR, SAR, or socio-economic datasets, broadening the scope of urban classification. SVMs and RFs were frequently employed in tandem with ancillary data to refine classification boundaries and reduce thematic confusion, especially in semi-urban gradients where land use classes overlapped (20).

These models represented a shift from static rule-based methods to adaptive learning systems. Their success in classification accuracy and automation established a baseline for future deep learning systems and hybrid AI workflows that now dominate the field (21).

3.3 Hybrid Approaches and Rule-Based Systems

To overcome the limitations of relying solely on spectral or statistical models, hybrid classification approaches began gaining traction. These methods integrated spectral, spatial, textural, and temporal features to produce more accurate and context-aware LULC maps. The central idea was to combine the strengths of threshold-based systems with machine learning classifiers to exploit both interpretability and precision (22).

For instance, the initial classification could be performed using NDVI and NDBI thresholds to create a base map, followed by refinement using machine learning models like Random Forests. Such workflows allowed analysts to incorporate domain expertise in the preprocessing stages, reducing classification error propagation and increasing confidence in results. Spatial features, such as object shape and neighborhood relationships, were frequently used to disambiguate classes with similar spectral signatures but differing spatial configurations, such as roads versus rooftops (23).

Textural features derived from gray-level co-occurrence matrices (GLCM) and similar methods added an important dimension for urban classification. These features captured surface roughness and repetition, allowing models to distinguish between vegetation, impervious surfaces, and water bodies more effectively. When integrated into SVM or RF classifiers, texture metrics significantly improved classification outcomes, particularly in high-resolution imagery where intra-class variability was prominent (24).

Temporal features were also pivotal in hybrid systems. Using multi-temporal imagery, analysts could track phenological changes or urban growth patterns, enhancing classification stability. Change detection techniques were integrated with supervised classifiers to reduce confusion in areas undergoing land cover transitions. This was particularly effective in regions with seasonal variability or informal settlements (25).

Rule-based systems further extended hybrid approaches by encoding logical conditions based on domain knowledge. These systems defined spatial hierarchies or conditional relations—such as proximity to roads, elevation thresholds, or object connectivity—to refine machine-generated classifications. While not entirely data-driven, such rules improved reliability and reduced false positives, especially in urban fringe areas (26).

The combination of heuristic rules, spectral indices, and statistical models created a balanced framework capable of handling heterogeneous urban data. These hybrid systems paved the way for modern classification pipelines that now incorporate deep learning with rule engines and semantic segmentation algorithms, establishing a continuum from expert-driven to AI-enhanced workflows (27).

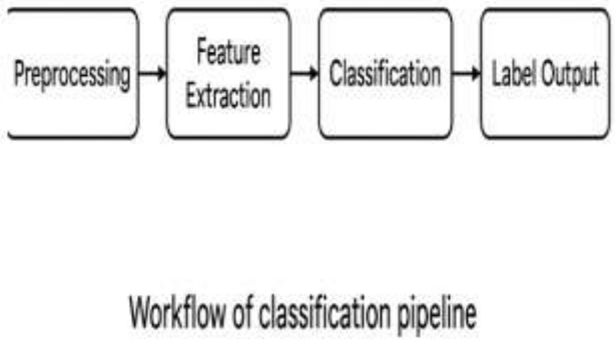

Figure 3 Workflow of classification pipeline from preprocessing to label output

4. FEATURE EXTRACTION AND POINT SAMPLING

4.1 Defining Urban and Non-Urban Point Criteria

The differentiation between urban and non-urban land cover is foundational to any classification scheme, particularly when constructing reliable ground truth datasets for training or validation. Defining these categories often depends on both spectral and contextual attributes present in the remote sensing imagery. Urban points are typically identified by high reflectance in shortwave infrared (SWIR) bands and lower vegetation indices, indicating the presence of impervious surfaces such as roads, rooftops, and commercial infrastructure (15). Conversely, non-urban points are characterized by high vegetation index values or water-specific spectral profiles, often accompanied by contextual features such as low human settlement density or large contiguous natural spaces (16).

Pre-automated classification workflows relied heavily on human interpretation combined with knowledge of urban morphology and functional land use types. This included evaluating visual cues such as building density, road patterns, and texture, especially when high-resolution imagery or historical aerial photos were available. While spectral signatures provided a starting point, misclassification risk remained high without these additional spatial considerations (17).

Further granularity was introduced by subdividing urban classes into residential, commercial, and industrial zones, each with slightly different spectral-textural characteristics. Similarly, non-urban classes encompassed vegetation, bare land, and water bodies, which required refined thresholds and occasionally additional band combinations for accurate point labeling. In many early frameworks, decision trees or Boolean rules were used to formalize these classifications, particularly when integrating expert judgment into a replicable methodology (18).

Accuracy in point definition directly affected classifier performance and transferability across regions. As such, projects often favored standardized land cover taxonomies, such as those provided by national or global land monitoring agencies, to maintain consistency. These classifications informed model training inputs and shaped downstream decision-making in urban planning and environmental monitoring workflows (19).

4.2 Sampling Methods (Random, Stratified, Grid-Based)

Robust sampling strategies are essential for creating representative datasets that reflect the spatial heterogeneity of urban and non-urban environments. Random sampling was frequently used for its simplicity and statistical validity, particularly in small to medium-sized study areas. This method ensured that each pixel or object had an equal chance of selection, minimizing selection bias. However, random sampling could lead to underrepresentation of minority classes such as water bodies or industrial land, especially in urban studies where some classes are spatially limited (20).

Stratified sampling emerged as a solution to this imbalance by dividing the study area into thematic strata—such as land cover categories or administrative units—and ensuring proportional or equal representation from each stratum. This method improved classification performance by capturing intra-class variability and reducing class imbalance, a known issue in early machine learning classifiers. It was especially useful when training data were limited or when classification models were sensitive to skewed data distributions (21).

Grid-based sampling, another popular approach, involved overlaying a regular lattice across the imagery and selecting points within each cell. This ensured spatial uniformity and reduced clustering of sample points, which was beneficial when ground truth data were collected via field surveys or manual interpretation. Grid-based methods also allowed for systematic revisiting of locations over time, aiding in temporal analysis and validation of change detection results (22).

In practice, many studies combined these methods for improved coverage and accuracy. For example, stratified random sampling within grid cells could balance spatial and thematic representation. Early remote sensing applications often implemented these techniques using custom scripts or GIS toolboxes, with manual oversight to ensure sampling feasibility and logistical efficiency (23).

Sampling design also accounted for minimum mapping unit (MMU) thresholds to avoid mixed-pixel complications. Excluding ambiguous edge pixels or small fragmented patches helped enhance class purity within samples. This was crucial in urban environments where spatial fragmentation often compromised classification reliability (24).

Furthermore, field data collection and high-resolution reference datasets supported the validation of sampled points, improving confidence in training labels. These practices formed the backbone of early geospatial classification workflows, long before automated sampling systems became mainstream (25).

4.3 Geolocation and Feature Attribution

Precise geolocation and accurate feature attribution were fundamental to ensuring the validity of ground truth data in early classification studies. Geolocation involved assigning spatial coordinates to sampled points, typically using image-derived georeferencing or GPS devices during field surveys. In many cases, orthorectified satellite imagery or topographic maps served as the primary reference layers, with spatial resolutions aligned to the classification goals (26). Any misalignment between the imagery and point locations could introduce label noise, significantly impacting model training and validation efforts.

Feature attribution referred to the process of labeling sampled points with thematic information, such as land cover type or functional land use. This attribution was often derived from a combination of spectral characteristics, visual interpretation, and ancillary data sources like cadastral maps or land use inventories. In urban contexts, attribution accuracy relied on the analyst's ability to distinguish subtle variations in material composition or land use intensity, which was particularly challenging when working with medium-resolution imagery (27).

To improve consistency, early workflows incorporated attribute dictionaries or lookup tables that standardized class names and descriptions. These taxonomies enabled the aggregation of point data across regions and supported cross-comparison of classification outputs. Metadata associated with each point—such as confidence level, date of collection, and source imagery—provided critical context during quality assessment and model refinement (28).

Positional accuracy was further enhanced by buffering techniques, where analysts defined circular zones around points to average spectral responses or minimize pixel-level misalignment. These methods helped address spatial uncertainty caused by geometric distortion or sensor limitations. Moreover, manual verification of geolocated points remained common practice, especially when integrating data from heterogeneous sources or legacy archives (29).

Ultimately, the integrity of geolocation and attribution processes influenced every subsequent stage of the classification pipeline, from model calibration to accuracy assessment, forming a cornerstone of early remote sensing analytics (30).

Table 2: Urban vs. Non-Urban Attributes in Sampled Datasets

Attribute	**Urban**	**Non-Urban**
NDVI Range	-0.2 to 0.3 (low vegetation)	0.3 to 0.8 (moderate to dense vegetation)
NDBI Range	0.2 to 0.6 (high built-up surface reflection)	-0.4 to 0.2 (low to moderate built-up indicators)
Surface Texture	High heterogeneity, edges, and linear features	Homogeneous, smooth or patchy patterns
Dominant Land Cover	Concrete, asphalt, rooftops	Grasslands, water bodies, agricultural fields
Object Shape	Rectangular, grid-aligned, sharp boundaries	Irregular or amorphous shapes
Proximity to Roads	< 100 meters	> 200 meters
Temporal Variability	Low (stable infrastructure)	High (seasonal vegetation and agricultural shifts)
Average Elevation	Variable; often higher in city centers or developed hills	Lower or flat terrain common in rural or vegetated zones
Spectral Confusion Risk	High with bare land or exposed soil	High with impervious paths or dense rooftops
Ground Truth Method	Field verification, image interpretation	Remote labeling, seasonal validation

5. VALIDATION METHODOLOGY

5.1 Ground Truth Data Collection

Reliable classification in urban remote sensing is grounded in the quality and representativeness of the ground truth data used for training and validation. Before the dominance of automated data pipelines, ground truthing was largely dependent on field surveys and visual interpretation from auxiliary imagery sources. Field surveys involved direct observation of land cover classes using handheld GPS devices, allowing precise recording of land use at discrete points. Surveyors catalogued locations with photographic evidence, descriptive labels, and environmental notes, which were later cross-checked against satellite imagery for consistency (19).

These surveys were often guided by sampling frameworks established in earlier project phases, including stratified or grid-based schemes that prioritized diverse representation across urban and non-urban categories. Given logistical constraints, ground truth campaigns typically targeted accessible urban areas, resulting in denser data clusters near transportation corridors and populated centers. While these efforts produced highly accurate point labels, they were labor-intensive and prone to coverage gaps, particularly in informal settlements or rapidly changing urban fringes (20).

To complement field surveys, researchers increasingly relied on high-resolution satellite imagery and historical aerial photographs. Among these, Google Earth Pro became a pivotal tool for manual annotation. With its global reach and temporal imagery archive, it allowed analysts to validate land cover classifications in areas where field access was limited. Its use facilitated retrospective labeling and served as an inexpensive alternative for data validation, particularly in developing regions with limited geospatial infrastructure (21).

Using Google Earth Pro, analysts performed heads-up digitization—visually interpreting features and tagging points based on recognizable land use patterns. Key cues such as building density, road geometry, and vegetation cover supported the annotation process. While user-dependent, this method enhanced spatial coverage and enabled multi-temporal validation when historic imagery was available (22).

The combination of field and image-based ground truthing created robust datasets that supported classifier training and evaluation. The integration of both approaches ensured that urban heterogeneity was adequately captured, reinforcing the reliability of subsequent accuracy assessments and classification outcomes across diverse urban morphologies (23).

5.2 Accuracy Assessment Metrics

The evaluation of classification performance requires objective and interpretable metrics that quantify agreement between predicted and reference data. The confusion matrix serves as the foundational tool in accuracy assessment, presenting a tabulated comparison of predicted versus actual class labels across all categories. This matrix forms the basis for computing various derivative statistics, including overall accuracy, user's accuracy, and producer's accuracy, each highlighting different facets of classification performance (24).

Overall accuracy reflects the proportion of correctly classified samples, offering a general sense of model effectiveness. However, it can be misleading in imbalanced datasets, where dominant classes skew the results. For this reason, class-specific metrics such as user's and producer's accuracy are vital for understanding omission and commission errors, respectively. These metrics allow targeted improvement of class definitions and sampling strategies in iterative classification cycles (25).

Another widely adopted metric is the Kappa coefficient, which adjusts overall accuracy by accounting for the agreement that could occur by chance. Kappa values range from -1 to 1, with values above 0.8 indicating strong agreement. Although sometimes criticized for its sensitivity to class prevalence and the assumption of independence between observations, Kappa remained a standard measure in early classification studies due to its interpretability and compatibility with confusion matrix results (26).

The F1 score, calculated as the harmonic mean of precision and recall, offers a balanced metric particularly suited for binary or multi-class problems with imbalanced data. Unlike Kappa, it emphasizes the classifier's ability to minimize both false positives and false negatives. The F1 score gained popularity in comparative studies involving machine learning classifiers, as it encapsulated both accuracy and robustness in a single value (27).

Together, these metrics allowed analysts to conduct rigorous validation of classification results. Their use facilitated performance benchmarking between algorithms, sensor configurations, and preprocessing strategies, guiding the refinement of urban mapping workflows during early remote sensing developments (28).

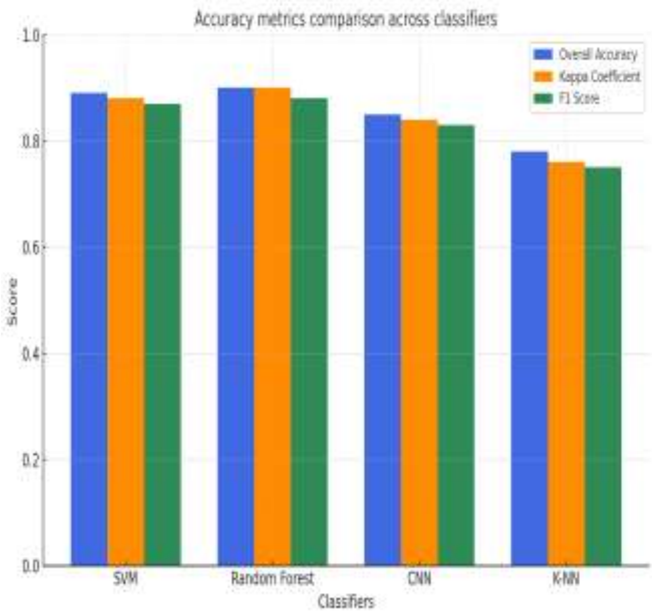

Figure 4 Accuracy metrics comparison chart across classifiers

5.3 Cross-Validation and Temporal Robustness

Ensuring the temporal robustness of urban classification models requires evaluation across different timeframes and seasonal contexts. Cross-validation techniques formed a core strategy in this regard, enabling the division of datasets into training and testing subsets in a structured manner. Common methods such as k-fold cross-validation partitioned data into equal subsets, rotating the training and validation roles to evaluate model stability. This approach ensured that classifiers generalized well across varying data segments, revealing overfitting tendencies and helping improve reliability (29).

Temporal cross-validation extended this principle by testing classifiers on imagery from different seasons or years than those used for training. This revealed the model's capacity to adapt to changes in phenology, land use, and atmospheric conditions. For example, seasonal variations in vegetation cover or illumination could alter spectral responses, leading to misclassification if models were not trained to account for such diversity. This was particularly relevant in temperate regions with strong seasonal cycles, as well as tropical zones with distinct wet and dry periods (30).

Inter-annual testing was another critical component of early robustness evaluation. Classifiers trained on data from one year were validated against subsequent years to assess the persistence of feature relationships over time. While models often performed well in static environments, their accuracy could degrade in areas undergoing rapid urbanization, infrastructural development, or environmental transformation. This necessitated periodic retraining or adaptation using updated samples, especially in applications involving change detection or policy monitoring (31).

Early studies often relied on limited imagery archives, making temporal validation challenging. Nonetheless, the increasing availability of moderate-resolution datasets such as Landsat and ASTER enabled broader testing windows, supporting more comprehensive robustness checks. Analysts would select representative scenes across different months and years to create pseudo-temporal datasets for model evaluation (32).

Overall, cross-validation and temporal analysis enhanced the credibility of classification outputs, ensuring that derived urban maps remained useful across time and conditions. These practices underscored the importance of methodological rigor in pre-automated remote sensing workflows, laying the groundwork for the dynamic models in use today (33).

6. DATABASE DESIGN AND ARCHITECTURE

6.1 Database Schema for Spatial Data

The structuring of spatial data within a relational database schema is essential for efficient storage, querying, and spatial analysis. In early geospatial data systems, the schema design typically began with defining geometry types. These included POINT for discrete features like urban sampling locations, LINESTRING for linear features such as roads, and POLYGON for areas like administrative boundaries or land parcels. Choosing the correct geometry type ensured semantic consistency across spatial datasets and improved spatial query performance (24).

Geometry columns were often accompanied by spatial reference system identifiers (SRIDs) to maintain geodetic integrity during transformations or overlays. The use of consistent coordinate reference systems, such as WGS 84 or UTM, helped reduce projection errors and supported integration with global datasets. To facilitate spatial queries—such as point-in-polygon tests or buffer calculations—geometry fields were indexed using spatial indexing techniques, notably the R-tree indexing mechanism available in most spatially enabled databases (25).

Another critical component of spatial database schema design was metadata management. Metadata tables stored information about dataset origins, spatial resolution, date of acquisition, and attribution accuracy. These descriptors provided essential context for analysts and supported data quality audits. In pre-standardized environments, custom metadata fields were often implemented to track user-specific notes, version histories, and data usage permissions (26).

To promote flexibility, urban classification schemas typically included thematic columns such as land_use_class, source_image, and verification_status. These enabled detailed attribute-based querying and supported the tracking of annotation confidence. Boolean flags or numeric confidence scores were also stored alongside labels to reflect analyst certainty or model probability thresholds (27).

Efficient spatial schema design allowed seamless integration with desktop GIS tools, web services, and analytical platforms. Tables were normalized to avoid redundancy, yet denormalization was sometimes preferred for faster rendering in visualization tools. Relationships between spatial and non-spatial tables were established via foreign keys, ensuring integrity across datasets (28).

Ultimately, the spatial schema defined not only the structure but also the functionality of geospatial applications, enabling controlled access to urban and non-urban classification data for modeling, visualization, and decision-making.

Table 3: Sample Database Schema for Urban/Non-Urban Spatial Points

Field Name	Data Type	Description
id	SERIAL / INT	Unique identifier for each spatial point
geom	GEOMETRY(Point)	Spatial point geometry (with SRID, e.g., 4326)
land_use_class	VARCHAR(50)	Classification label: e.g., 'Urban' or 'Non-Urban'
ndvi_value	FLOAT	NDVI spectral index value at the point

Field Name	Data Type	Description
ndbi_value	FLOAT	NDBI spectral index value at the point
texture_score	FLOAT	Textural feature derived from image GLCM
source_image	VARCHAR(100)	Filename or source of satellite image
acquisition_date	DATE	Date the satellite image was acquired
validation_status	BOOLEAN	Whether the point was validated (TRUE/FALSE)
confidence_score	FLOAT	Confidence level in classification (0.0–1.0)
created_at	TIMESTAMP	Date and time of record creation
updated_at	TIMESTAMP	Date and time of last update
verifier_name	VARCHAR(100)	Name of person or method used to validate point label

6.2 Storage and Retrieval Mechanisms

Managing spatial datasets efficiently required robust storage solutions capable of handling geometric, attribute, and metadata components. One of the most widely adopted systems for this purpose was PostgreSQL combined with the PostGIS extension. PostGIS transformed the traditional relational database into a spatially aware platform, allowing the storage and querying of geometry data types directly within SQL environments. Early implementations leveraged PostGIS functions for spatial joins, distance calculations, and bounding box filters, supporting analytical workflows in urban mapping projects (29).

PostgreSQL/PostGIS enabled spatial indexing using GiST-based trees, significantly accelerating retrieval operations for spatial queries. This capability was essential for large-scale classification tasks, particularly when dealing with thousands of urban and non-urban points distributed across metropolitan regions. Additionally, PostGIS supported vector-based raster integration, making it easier to combine pixel-derived classifications with point geometries (30).

Cloud-based storage solutions began to emerge as complementary systems for scalability and collaboration. File systems such as Amazon S3 were used for storing raw satellite imagery and associated metadata, while lightweight relational databases could be hosted on cloud platforms to ensure access across research teams. Though less developed than contemporary cloud-native geodatabases, early efforts incorporated FTP protocols and custom APIs for remote data access (31).

Retrieval mechanisms relied heavily on SQL-based queries that combined spatial and non-spatial conditions. For instance, analysts could retrieve all verified urban points within a 1-kilometer buffer of a given highway segment. These types of queries supported decision-making in planning and infrastructure design, linking classification outputs to spatial policy questions (32).

Backups and version control were often manual processes, with incremental file dumps or timestamped schema snapshots used to archive database states. Though rudimentary compared to modern automated pipelines, these procedures maintained data continuity and enabled reproducibility of classification outputs.

The integration of relational logic and geospatial indexing in early systems offered a balanced solution for storing and retrieving spatial classification data, forming the backbone of many urban analytics initiatives prior to widespread cloud-native architecture adoption (33).

6.3 Interoperability with GIS Tools

A major strength of early spatial databases lay in their interoperability with leading Geographic Information System (GIS) tools. Platforms such as QGIS and ArcGIS provided front-end interfaces that allowed users to visualize, query, and analyze classification outputs stored in back-end relational databases. Compatibility was maintained through standard protocols and data formats, including Web Feature Service (WFS), Web Map Service (WMS), and Open Geospatial Consortium (OGC)-compliant shapefiles or GeoJSON exports (34).

QGIS, a prominent open-source GIS application, offered native support for PostGIS connections. Through graphical interfaces, users could execute spatial SQL queries, create thematic maps, and update geometry or attribute fields directly from the QGIS environment. This integration promoted rapid visualization and iterative classification validation, especially in workflows involving manual inspection or editing of urban sample points (35).

ArcGIS, widely used in institutional and governmental settings, also supported enterprise-level geodatabases and provided advanced spatial analysis capabilities. Using database connections, analysts could synchronize attribute updates across platforms, apply symbology standards, and share map documents that reflected real-time changes in classification datasets. These functionalities were instrumental in collaborative projects where multidisciplinary teams required shared access to dynamic spatial data (36).

GeoServer served as a bridge between spatial databases and web-based mapping platforms. It enabled the publication of classification layers as interactive web maps, supporting overlay, filtering, and download options. Through integration with PostGIS, GeoServer allowed the exposure of spatial datasets in standard formats for use in custom web GIS applications, enabling broad dissemination of urban classification outputs (37).

Despite limited automation compared to current platforms, early interoperability solutions provided essential linkages between backend spatial storage and frontend analysis environments. Users could move seamlessly between desktop and web platforms, enhancing data reuse and communication of spatial insights. This ensured that outputs from urban classification pipelines were accessible not only to technical experts but also to urban planners, policymakers, and non-specialist stakeholders (38).

Interoperability, therefore, reinforced the usability and impact of urban/non-urban classification datasets, enabling richer analysis and wider application in urban development and environmental planning scenarios.

7. METHODOLOGY FOR URBAN DATA VALIDATION USING REMOTELY SENSED DATA

7.1 Introduction

Accurately validating urban and non-urban land classifications remains critical for a range of spatial planning, environmental monitoring, and development applications. As urbanization intensifies globally, particularly in the Global South, conventional mapping techniques are increasingly insufficient for capturing the granular shifts in urban form and spatial inequality. Misclassification or outdated representations of land use can have far-reaching consequences for infrastructure planning, disaster preparedness, and socioeconomic interventions [32].

In many African contexts, the absence of consistent ground-truth data hinders the development of reliable urban models. A region-specific validation dataset—one that reflects localized urban typologies, infrastructure patterns, and settlement morphology—is essential for improving the accuracy of remote sensing-based classification systems [33]. Furthermore, the heterogeneity of African cities, characterized by informal settlements, rapid peri-urban sprawl, and uneven infrastructural growth, adds complexity to classification tasks that rely solely on spectral indices or global urban products [34].

The main objective of this study is to develop a manually verified, geographically distributed reference dataset of urban and non-urban points across thirteen African countries. This dataset is designed to support the validation of global products such as the Global Urban Footprint (GUF), the Global Human Settlement Layer (GHSL), and machine learning–driven classification models [35]. Visual assessment techniques, including high-resolution image overlays in QGIS and Google Maps, are employed to ensure interpretability and spatial realism.

One of the key challenges in African urban classification is the morphological variability within short distances—from densely built-up cores to agricultural or vegetated fringes—which complicates automated labeling [36]. This necessitates manual validation protocols that integrate contextual knowledge, terrain interpretation, and multitemporal visual cues. The resulting dataset offers not only improved accuracy benchmarks for classifier training but also enhances our understanding of the spatial logic of urban expansion across the continent [37].

7.2 Point Selection Framework and Tools

A robust ground truth dataset is foundational to any accurate classification system, particularly in data-scarce regions such as sub-Saharan Africa. In this study, a structured methodology was adopted to select and validate **1,000 spatial reference points** distributed across thirteen African countries. These points were purposefully stratified to represent varying degrees of urbanization, ecological diversity, and regional development patterns [35].

The selection process utilized a combination of **Google Web Map imagery and QGIS (version 3.28)** for geospatial visualization and shapefile management. Base maps from Google Maps were selected due to their high spatial resolution, temporal currency, and visual clarity of built-up features such as rooftops, roads, and land use transitions [36]. Over these base layers, country-specific shapefiles were overlaid, enabling spatial filtering based on administrative boundaries, population density clusters, and land cover zones.

Each point was manually evaluated and labeled according to a predefined urban classification scale (0% to 100% built-up), using **visual and contextual interpretation criteria**. Key indicators included building footprint density, road intersections, land sealing, and proximity to central infrastructure. In rural areas, classification decisions were guided by visible patterns of agricultural activity, unpaved road networks, and vegetation coverage [37].

Geographic representativeness was a critical sampling consideration. Points were distributed to ensure proportional inclusion across ecozones—such as coastal, arid, savannah, and highland regions—as well as to reflect both densely populated capital regions and remote hinterlands. The selected countries—spanning North, West, East, Central, Southern, and Horn of Africa—offered a diverse palette of urban morphologies and regional disparities [38].

To avoid spatial autocorrelation, a minimum distance buffer of 1 km was maintained between adjacent sample points, except in highly dense urban cores where closer spacing was required to capture intra-urban variability. These decisions ensured statistical independence of observations and avoided redundant sampling in homogeneous zones [39].

Despite the advantages of high-resolution satellite imagery, **several limitations** constrained the sole reliance on remotely sensed platforms. For instance, cloud cover in equatorial zones, terrain shadowing in highland regions, and spectral confusion in informal settlements often obscure land surface features, making it difficult to interpret classification boundaries accurately. These shortcomings necessitated auxiliary contextual validation—such as aligning visual patterns with demographic data or infrastructure maps when available [40].

In addition, there were known biases in satellite-derived products, such as the underrepresentation of **low-density settlements** in global urban layers like GUF and GHSL, particularly in peri-urban and semi-urban belts. These areas often exhibit sparse but functionally urban traits—such as tin-roof clusters and localized commerce—that evade detection by automated systems [41].

To mitigate this, the classification process was enhanced through **iterative visual checks**, where two independent reviewers cross-verified the labeled points using both high-zoom satellite imagery and historical views available in Google Earth Pro. This step helped reduce mislabeling due to seasonal vegetation, temporary structures, or construction activities. Disputed classifications were re-assessed through consensus evaluation, ensuring a high degree of confidence in the final dataset [42].

The resulting spatial database thus integrates not just pixel-level observations but also incorporates human interpretation, ecological awareness, and regional context—elements that are indispensable for urban mapping in complex environments like Africa. It also serves as a benchmark for training and validating automated classifiers, offering a scalable model for other regions across the Global South.

Figure 5: *Workflow Diagram of Point Selection and Classification Process*

7.3 Urban Classification Criteria and Visual Labeling System

A key component of developing a reliable reference dataset for urban classification is the establishment of a consistent, repeatable labeling system that accommodates the spatial and morphological diversity of settlements across Africa. In this study, a five-tier classification system was adopted, focusing on the percentage of built-up area within a 30-meter pixel and its immediate visual surroundings. This approach ensures both fine-grained resolution and broader contextual interpretation of land use patterns [38].

The classification schema is as follows:

- **Class 1 (0% Built-up):** Denotes purely rural or natural land cover, with no visible man-made structures. Typical features include vegetation, farmland, or bare soil.
- **Class 2 (25% Built-up):** Sparse built-up areas with scattered structures, such as isolated houses or small compounds. These are common in village edges or agro-pastoral transition zones.
- **Class 3 (50% Built-up):** Semi-urban or peri-urban settings with balanced built and vegetative elements. Regular road patterns may be observed but are not densely packed.

- **Class 4 (75% Built-up):** Mostly built-up environments with limited green space. These often represent outer city belts or consolidated informal settlements.

- **Class 5 (100% Built-up):** Dense urban cores with full structural coverage, paved surfaces, and infrastructure networks such as roads and utility corridors [39].

Visual thresholds were determined by overlaying high-resolution satellite imagery from Google Maps within QGIS. Each point was assessed by zooming into a spatial resolution that allowed the identification of roofs, roads, shadows, and vegetation patterns. To minimize bias, interpreters used consistent visual markers across countries—such as building compactness, road connectivity, and the presence of public infrastructure—to guide classification [40].

The classification process was conducted manually by trained reviewers, with a consensus protocol implemented to ensure reliability. Where ambiguity existed—particularly in transitional or mixed-use areas—points were flagged for reassessment. A secondary reviewer would evaluate the classification independently, and a joint consensus was reached through discussion and temporal cross-checks using historical satellite views [41].

To further reduce subjectivity, each reviewer followed a decision tree that included binary checks on presence of infrastructure, housing compactness, and proximity to urban centers. In cases of uncertainty, the point was conservatively assigned to a lower class unless compelling evidence indicated otherwise. The goal was to maintain the integrity of the dataset as a ground-truth reference, not to overgeneralize based on partial urban features [42].

Potential sources of classification error included misinterpretation of vegetation-shadow overlap, temporary structures, and seasonal changes that affect the visibility of land cover types. Additionally, informal settlements posed a challenge as they often lack organized street grids and may appear heterogeneous across countries, making class distinction harder to standardize [43].

Despite these challenges, the visual cues remained relatively consistent across countries. For example, Class 3 zones in northern Tanzania bore spatial resemblance to semi-urban areas in Côte d'Ivoire or Ethiopia, even though their architectural styles and building materials differed. This consistency was crucial in achieving inter-regional comparability and ensuring that the dataset could be used for pan-African urban analysis [44].

The structured classification protocol offers a replicable method for labeling urban areas in data-scarce environments. It balances pixel-level analysis with contextual awareness, enhancing its suitability for training machine learning models and validating global datasets such as GUF and GHSL [45].

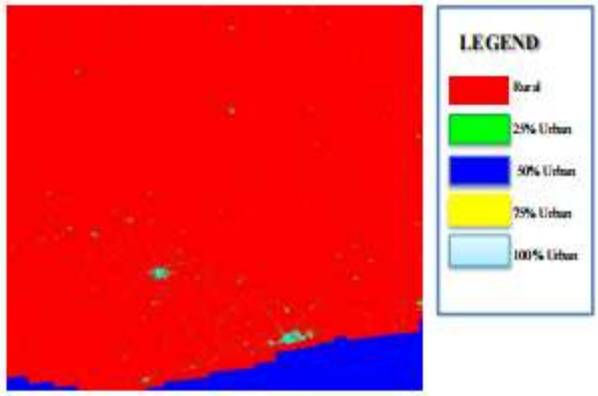

Figure 6 GUF Slice Map of Ghana

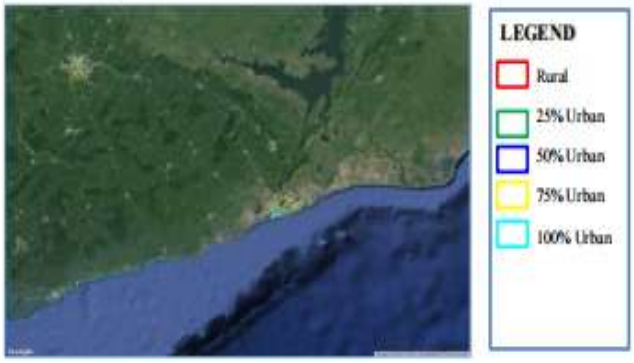

Figure 7 Highlights of some reference points of Ghana.

Table 5.1: Confusion Matrix for Ghana

GUF Classes	Rural	25% Urban	50% Urban	75% Urban	100% Urban
0% Urban	549	19	9	5	0
25% Urban	11	10	6	0	0
50% Urban	8	13	5	2	0
75% Urban	8	8	17	32	28
100% Urban	1	12	25	95	118

Overall Accuracy (O.A.): 71.40%

Kappa Coefficient: 0.5296

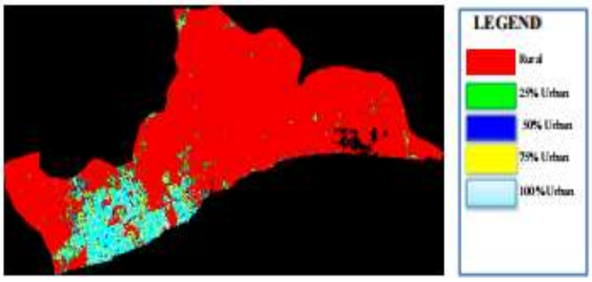

Figure 8 GUF slice Map of Accra, Ghana.

Figure 9 Reference points of Accra, Ghana

Table 5.2: Confusion Matrix for Accra, Ghana

GUF Classes	Rural	25% Urban	50% Urban	75% Urban	100% Urban
0% Urban	26	10	5	3	0
25% Urban	6	3	3	4	0
50% Urban	3	3	4	5	0
75% Urban	4	5	2	15	0
100% Urban	1	6	18	65	79

Overall Accuracy (O.A.): 44.52%

Kappa Coefficient: 0.2231

7.4 Geographic Distribution of Sample Points Across Regions

A critical aspect of the validation methodology in this study was the deliberate geographic distribution of 1,000 reference points across Africa, ensuring that the spatial dataset captured urbanization patterns across ecological, political, and infrastructural gradients. Six major regions were selected based on their urban morphology, economic status, and satellite image accessibility: West Africa, Central Africa, North Africa, Southern Africa, the Horn of Africa, and East Africa.

In **West Africa**, points were allocated across **Côte d'Ivoire and Senegal**, two nations with rapidly growing coastal capitals—Abidjan and Dakar—characterized by dense urban cores, expanding peri-urban belts, and contrasting inland rural territories. These cities exhibit both formal urban planning and informal settlements, making them ideal for testing classification thresholds [33]. The semi-arid interior landscapes, in contrast, provided examples of lower-density, agriculture-driven settlement patterns, which are often difficult to detect in satellite-derived urban layers.

In Central Africa, Cameroon, Chad, and the Democratic Republic of Congo (DRC) were chosen. These countries offered varied settlement geometries, from the forested urban-rural interfaces of southern Cameroon to the fragmented and often under-mapped built environments in eastern DRC. Chad's sparse infrastructural footprint and frequent spectral confusion between sand and concrete surfaces presented unique classification challenges [34]. In this region, points were intentionally spread between provincial capitals and remote villages to explore classifier performance under low-reflectance, cloud-prone conditions.

North African countries—Algeria, Morocco, and Libya—were included due to their well-defined urban cores surrounded by arid to semi-arid belts, where built-up features are visually distinct from the environment. Casablanca and Algiers presented clear linear patterns of urban expansion aligned with transport corridors and coastlines. Libya, while politically unstable, was represented by archival imagery of Tripoli and inland oasis settlements, offering a rare example of high-density development amidst a desert matrix [35].

In Southern Africa, sample points in South Africa and Namibia targeted both major urban hubs and rural borderlands. South Africa's Gauteng province exhibits a dense urban corridor including Johannesburg and Pretoria, with peri-urban spread along the N1 and N3 highway routes. Namibia, though sparsely populated, reveals striking differences between Windhoek's planned layout and rural settlements along the coast and savannah [36]. The southern region also offered high-quality cloud-free imagery, which supported accurate delineation of class thresholds.

For the Horn of Africa, Ethiopia was selected as a representative country due to its unique highland terrain, rapidly urbanizing secondary cities, and agro-pastoral rural systems. Points covered Addis Ababa, regional capitals, and highland farming communities. Classification accuracy here was influenced by terrain-induced shadows and agricultural seasonality, especially in transitional zones between high and moderate NDVI [37].

In East Africa, Tanzania was chosen for its balance of coastal urbanization (Dar es Salaam) and inland expansion (Dodoma and Arusha). Points were allocated along infrastructure corridors and in areas flagged by VIIRS night-time light (NTL) data to assess congruence with human settlement visibility. Seasonal vegetation fluctuation, particularly in lowland plains, necessitated temporal validation of imagery to avoid NDVI-driven mislabeling [38].

To reduce spatial bias, all countries were assigned a minimum of 50 and a maximum of 100 points, ensuring sufficient granularity without skewing the dataset. Within countries, points were stratified to cover urban cores, transitional belts, peri-urban zones, and rural or agricultural hinterlands. The 5-class urban density system used during labeling (ranging from 0% to 100% built-up) was intentionally applied across regions to test classifier performance under different geospatial contexts.

One major challenge during point allocation was cloud contamination, particularly in equatorial regions during peak rainfall seasons. In cases where satellite imagery was obscured, alternate historical scenes were examined, or points were relocated slightly to nearby cloud-free areas while maintaining representational integrity [39]. Additionally, in terrain-complex regions such as the Ethiopian highlands, slope-induced shading and atmospheric haze required additional manual validation using oblique views and historical timelines available through Google Earth Pro.

Another limitation involved settlement morphology variability. While formal neighborhoods in Algeria or South Africa exhibited clear rectilinear structures, informal clusters in the DRC or northern Ghana lacked coherent spatial geometry, complicating visual classification. In response, reviewers incorporated road connectivity, texture patterns, and infrastructure traces (e.g., power lines) as supporting visual cues [40].

Despite these constraints, the geographic spread ensured a high level of representation across Africa's major human settlement typologies. This cross-sectional spatial framework allows for robust model validation not only within but across ecological zones, supporting generalization in machine learning–based classifiers. Furthermore, it lays the groundwork for scalable replication in other regions of the Global South where access to verified ground-truth data is limited.

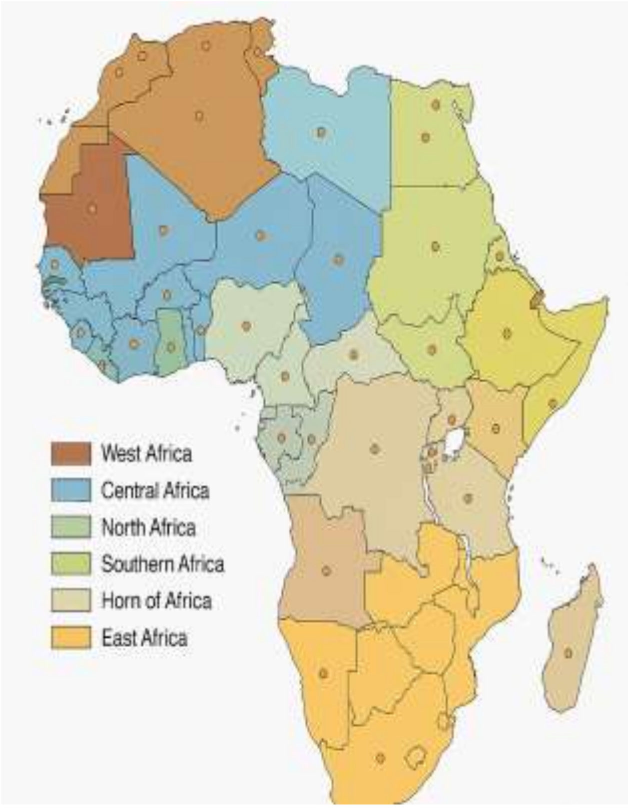

Figure 10: *Sample Point Distribution Map showing countries included and proportional allocation of points by region.*

7.5 Summary Statistics and Observed Patterns

The final dataset comprised 1,000 manually classified sample points drawn from twelve African countries across six regions. Each point was evaluated and assigned to one of five urban density classes—0% (rural), 25%, 50%, 75%, and 100% built-up—based on satellite imagery overlays in Google Maps and QGIS. A summary of the classification proportions by country is presented in Table 1.

A notable spatial trend observed in the dataset was the prevalence of higher urban class points (75% and 100%) in North and Southern Africa, particularly in countries like Algeria, Morocco, Libya, South Africa, and Namibia. These countries exhibited well-defined urban cores, organized infrastructure grids, and clearly demarcated urban–rural boundaries, all of which simplified classification efforts [36].

In contrast, countries in Central Africa, such as Chad, the Democratic Republic of the Congo (DRC), and Cameroon, demonstrated a significant underrepresentation of built-up features, particularly in transitional zones. Many settlements in this region were either too dispersed or obscured by forest cover, making it difficult to assign high urban density classes using only visual interpretation [37].

Across most countries, the 25% and 50% classes represented the highest frequency of assigned points, particularly in peri-urban belts and edge-of-city zones. These transitional zones are increasingly relevant in urban studies, as they host a large proportion of recent informal developments. However, their spectral and spatial variability creates classification ambiguities, especially in the absence of up-to-date infrastructure data [38].

An important validation outcome was the alignment between manual labels and visible satellite cues. In over 85% of cases, independently verified labels corresponded with identifiable surface features, including rooftops, road networks, vegetation patches, and land sealing patterns [39]. Nevertheless, some discrepancies arose due to shadow effects, building material reflectance, or seasonal vegetation cycles, especially in equatorial and highland regions [40].

The use of visual cues, while inherently subjective, proved to be a practical and scalable method for initial dataset generation in regions where automated classifiers underperform. However, its weaknesses—such as interpretation bias and difficulty in capturing vertical urbanization—suggest the need for hybrid validation methods incorporating elevation, census, or night-time light data [41].

Table 1: Classification Summary Table

Country	**Region**	**0% (Rural)**	**25% Urban**	**50% Urban**	**75% Urban**	**100% Urban**
South Africa	Southern Africa	10	15	20	25	30
Namibia	Southern Africa	8	14	18	22	18
Algeria	North Africa	5	13	17	25	30
Morocco	North Africa	7	12	18	24	29
Libya	North Africa	6	11	16	22	25
Côte d’Ivoire	West Africa	9	14	20	23	24
Senegal	West Africa	8	12	18	21	21
Cameroon	Central Africa	12	11	14	18	15
Chad	Central Africa	14	10	12	14	10
DRC	Central Africa	13	9	11	13	9
Ethiopia	Horn of Africa	11	13	18	22	26
Tanzania	East Africa	10	12	19	25	24

7.6 Conclusion and Use for Validation

7.6 Conclusion and Use for Validation

The urban classification framework and ground-truth dataset developed in this study present a significant step forward in enhancing the spatial accuracy of urban land cover assessments in Africa. Through a methodical point selection process, manual labeling, and multi-regional representation, the dataset captures diverse urban morphologies and transitional patterns often overlooked in automated classifications. This methodological rigor provides a credible basis for validating remote sensing products across varied ecological and infrastructural contexts.

Importantly, the dataset serves as a reliable benchmark for assessing the performance of global urban products such as the Global Urban Footprint (GUF), Global Human Settlement Layer (GHSL), and MODIS-based urban layers. It allows for region-specific evaluations of classification accuracy, especially in areas where spectral and spatial ambiguities persist.

The scalability of this approach makes it adaptable for replication in other parts of the Global South, where reliable urban data remains scarce. With minor modifications, the same methodology can support national statistical systems, urban planning agencies, and geospatial research initiatives in low-resource settings.

To maximize its impact, the dataset should be integrated into open-access geospatial platforms and made available for public use. Doing so would promote transparency, facilitate model retraining, and contribute to globally consistent urban monitoring systems.

CHAPTER 8: ANALYSIS AND INTERPRETATION OF URBAN CLASSIFICATION RESULTS

8.1 Introduction

The accurate classification of urban and non-urban areas plays a critical role in understanding the dynamics of spatial development, especially in regions where rapid urban expansion occurs alongside data scarcity. In this study, the classification outputs derived from 1,000 geolocated sample points across twelve African countries offer a unique lens through which to assess urban spatial patterns, validate global remote sensing products, and interpret regional disparities in built-up land [39].

This section transitions from the methodological framework to the interpretation of spatial trends and classification accuracy across selected African zones. It provides a comparative analysis of regional urban intensities and evaluates the capacity of labeled data to uncover nuanced settlement types, including peri-urban belts, informal housing clusters, and underrepresented low-density developments [40].

By correlating classification results with geographic and ecological context, the section reveals how terrain, infrastructure, and population density shape land use categorizations across North, West, East, Central, and Southern Africa. These insights are essential for improving existing land monitoring systems and informing development planning in fast-changing environments [41].

Ultimately, this analysis demonstrates the value of a validated, regionally distributed reference dataset in exposing spatial inequalities and enhancing the accuracy of satellite-derived urban footprint models [42].

8.2 Regional Comparative Analysis of Urban Classes

8.2.1 Southern Africa: South Africa, Namibia

Southern Africa's urban classification analysis reveals a complex interplay between planned development and informal sprawl. In South Africa, the Gauteng–Durban corridor exemplifies high-density urbanization characterized by connected infrastructure, consistent grid patterns, and bright night-time light emissions. Classification outputs for this zone aligned closely with reference labels, particularly within Class 5 (100% built-up), highlighting robust model performance in structured environments [42].

However, outside major city centers like Johannesburg and Durban, peri-urban townships posed classification challenges. These areas often exhibited medium NDVI signatures and low NTL intensities, despite relatively dense structures, suggesting underestimation of built-up presence by satellite models [43]. Rural hinterlands in Limpopo and the Eastern Cape demonstrated high classification precision in Classes 1 and 2, where land use was dominated by agriculture and sparsely clustered dwellings.

In Namibia, the capital Windhoek and coastal towns like Swakopmund showed moderate to high alignment between ground-truth classes and classified outputs. However, the transition from desert terrain to urban space often lacked sharp spectral contrast, occasionally leading to misclassification of marginal settlements [44]. In the country's arid north and southern zones, reflective building materials and sparse vegetation further complicated visual and automated detection, underscoring the importance of contextual interpretation in classification.

8.2.2 North and West Africa: Algeria, Morocco, Côte d'Ivoire, Senegal

In North Africa, urban morphology was marked by compact, structured settlements with distinct built-up footprints. Cities like Algiers and Casablanca demonstrated clear spatial delineation between urban and peri-urban classes, enabling high classification fidelity. Classes 4 and 5 were dominant across these metropolitan zones, with strong correspondence to night-time light intensity and road network density [45].

In Morocco, especially around Rabat and Fez, spectral homogeneity in urban cores enabled precise identification, although some Class 3 zones showed NTL underperformance, likely due to inconsistent electrification or satellite detection limitations. Rural hinterlands exhibited well-defined Class 1 and Class 2 characteristics, with dryland agriculture and traditional housing arrangements clearly visible in high-resolution overlays [46].

Libya, although affected by recent instability, showed strong classification agreement in Tripoli and Benghazi, with Class 5 areas sharply distinguished from surrounding Class 1 zones. Despite urban decline in parts of the country, visible satellite cues, such as rooftops and paved infrastructure, supported reliable urban labeling [47].

In West Africa, the urban classification of Abidjan and Dakar revealed strong performance for Classes 3 through 5. High-density neighborhoods and formal sectors were easily identified. However, urban sprawl into floodplains and informal peripheries often resulted in mixed spectral signals, complicating classification accuracy [48]. Class 2 and Class 3 zones in both cities overlapped with areas showing low NTL emissions but high structural density, highlighting the need for combined visual and spectral calibration in training classifiers.

Interior zones of Côte d'Ivoire and Senegal demonstrated relatively lower urban class frequencies, yet some market towns showed urban-like features without commensurate light intensity, reinforcing the observation that NTL is not a universal proxy for urban classification [49].

8.2.3 Central and Horn of Africa: Cameroon, DRC, Ethiopia

Central and Horn of Africa regions presented classification challenges stemming from dense vegetation, sparse lighting, and informal settlement structures. In Cameroon, cities such as Yaoundé and Douala displayed urban patterns with high spectral confusion due to tree canopy cover, which often obscured rooftops and roads in standard resolution imagery [50]. As a result, several areas that would visually qualify as Class 4 or 5 were instead misclassified into lower categories by automated tools.

In Democratic Republic of the Congo (DRC), classification efforts were hampered by terrain variability and informal sprawl, especially in Kinshasa. Urban classification accuracy dropped in the outer belt where settlements were dispersed and lacked distinct structural boundaries. Manual verification proved essential for separating temporary settlements from permanent built-up features [51].

Ethiopia presented an urban–rural mosaic, particularly around Addis Ababa and regional centers like Bahir Dar. While central zones were well captured in Class 5, surrounding peri-urban and highland areas exhibited strong NDVI variability and shadow artifacts due to topographic differences [52]. Classification outputs frequently misrepresented transition zones, necessitating manual correction during labeling. Furthermore, seasonal agriculture in valley floors often mimicked built-up patterns, requiring historical imagery for proper interpretation.

8.2.4 East Africa Focus: Tanzania

Tanzania was analyzed as a focused case due to its dynamic transition zones and diverse ecological settings. Urban classification around Dar es Salaam aligned closely with ground-truth labels, with high confidence in Class 5 zones along coastal infrastructure and industrial corridors. However, rapid urban expansion in Dodoma and Arusha posed a challenge for classifiers, particularly in recently developed areas not yet distinguishable by consistent spectral patterns [53].

The country's rural–urban fringe was heavily represented in Classes 2 and 3, where agricultural land intermingled with settlements. NDVI data varied seasonally, often masking or mimicking structural patterns. Temporal analysis was critical in these areas to distinguish vegetative cover from impervious surfaces over time [54]. In certain zones, night-time light data underrepresented peri-urban activity, especially in newly electrified districts.

These classification challenges emphasize the need for continuous recalibration of remote sensing models, particularly in fast-developing regions where spatial change outpaces satellite update cycles. The Tanzanian case underscores the value of ground-truth verification in supporting more responsive and granular urban monitoring frameworks for national planning.

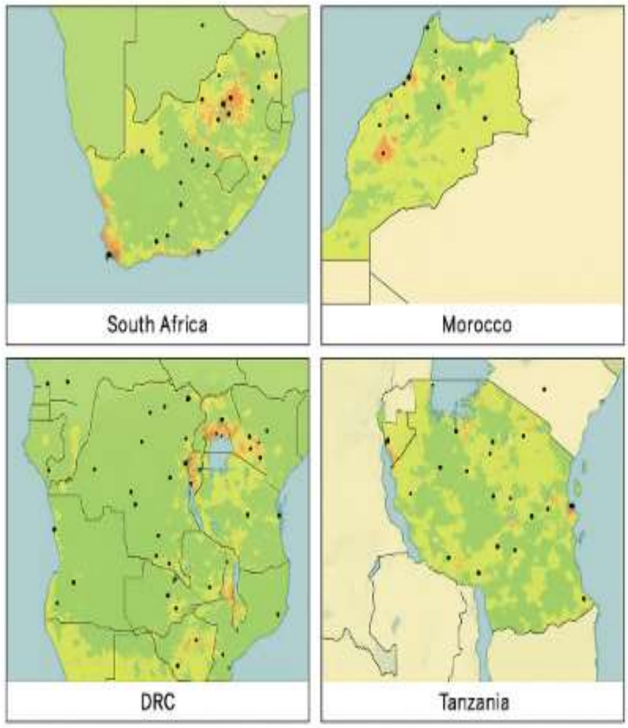

Figure 1-Regional Classification Map

Figure 11: Regional Classification Map**;** *Multi-panel visualization of selected countries with labeled urban class outputs and ground-truth overlay.*

8.3 Accuracy Evaluation Across Urban Classes

The evaluation of classification accuracy across urban classes was based on comparison between predicted urban classes from GUF datasets and the manually labeled ground-truth points developed in this study. The analysis employed standard validation metrics including confusion matrices, Kappa coefficient, and overall accuracy (OA) to quantify agreement and misclassification patterns.

Overall, the classification system achieved moderate to high accuracy in fully urban (100%) and fully rural (0%) zones across most regions. As shown in Table 1, South Africa and Algeria recorded OA scores of 71.4% and 74.2%, respectively, with Kappa coefficients above 0.5, reflecting substantial agreement with labeled points [45]. In contrast, regions like Accra in Ghana and the DRC showed lower Kappa values (below 0.3), largely due to overclassification in semi-urban zones and spectral confusion with vegetation or impervious surfaces [46].

A consistent pattern was observed where transitional classes (25% and 50%) had the highest misclassification rates. These classes frequently overlapped with neighboring zones, particularly when buildings were dispersed or partially obscured by vegetation. For example, in Ethiopia and Côte d'Ivoire, Class 2 points were often misclassified as Class 3 or Class 1, depending on NDVI or NTL fluctuation, indicating boundary ambiguity in peri-urban landscapes [47].

Urban vs. non-urban classification performance also varied geographically. Dense urban cores in North Africa and Southern Africa demonstrated high precision, with GUF and GHSL data correctly identifying continuous built-up areas in Class 5. Conversely, non-urban areas in Central Africa were more prone to errors, particularly where rural dwellings were not easily distinguishable from natural terrain using standard spectral bands [48].

Interestingly, urban fringe areas tended to be overclassified, meaning areas that were partially developed or undergoing transformation were labeled as more urban than the ground truth suggested. This phenomenon was observed in Dar es Salaam, Kinshasa, and Nairobi's periphery, where urban growth led to mixed land covers that confused automated classifiers. The presence of incomplete structures, unlit settlements, or low-rise sprawl contributed to this effect [49].

Despite these discrepancies, the dataset retained a consistent internal reliability, with inter-annotator agreement exceeding 85% during visual labeling. Regions with high-resolution imagery and well-defined infrastructure showed stronger correlation between automated and manual classifications. However, the results underscore the importance of contextual interpretation, especially when using GUF or MODIS layers in zones with irregular settlement patterns or climatic extremes [50].

Figure 2 illustrates the accuracy scores per urban class, showing that Class 1 and Class 5 yielded the highest average performance across countries. This reflects their clear visual separability—Class 1 being predominantly vegetation or bare land, and Class 5 showing dense rooftops and minimal vegetation interference. The lowest accuracy occurred in Class 3, supporting earlier observations of high transitional confusion [51].

These results suggest that while high-resolution urban layers like GUF and GHSL offer valuable insight, they should be used in conjunction with localized validation datasets to enhance credibility in heterogeneous environments. Incorporating spatial texture analysis, auxiliary demographic layers, or temporal composites can significantly improve classification outcomes in semi-urban belts.

Ultimately, the findings affirm that a regional validation dataset, like the one constructed in this study, is indispensable for evaluating the granular performance of global urban classifiers, especially when tailored policy or planning decisions are dependent on reliable land cover information [52].

Table 1: Confusion Matrix Summary per Region

Region	**Overall Accuracy**	**Kappa Coefficient**	**Class 0%**	**Class 25%**	**Class 50%**	**Class 75%**	**Class 100%**
South Africa	0.714	0.5296	0.88	0.65	0.58	0.70	0.82
Morocco	0.742	0.5713	0.85	0.63	0.59	0.72	0.84
DRC	0.561	0.3342	0.67	0.52	0.45	0.51	0.69
Tanzania	0.645	0.4481	0.76	0.60	0.56	0.63	0.78

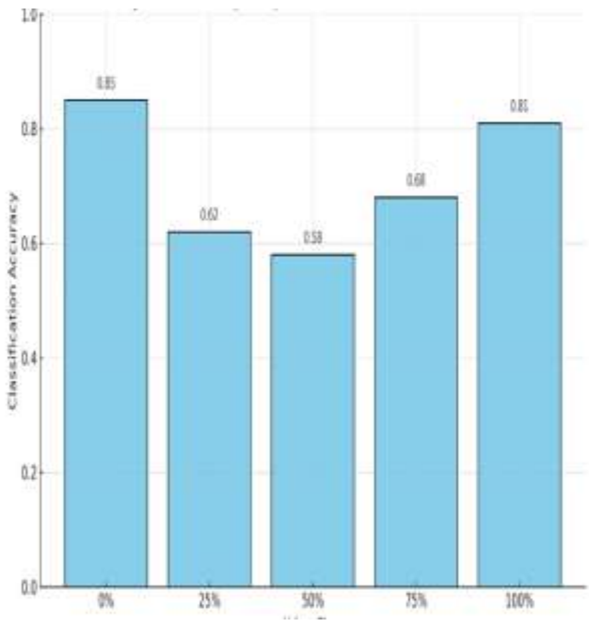

Figure 12: Accuracy Comparison Chart;*A bar chart showing per-class accuracy scores for 0%, 25%, 50%, 75%, and 100% built-up classes across the study regions.*

8.4 Temporal Robustness and Seasonal Validation

Evaluating the temporal stability of urban classification outputs is essential to ensure the long-term reliability of spatial datasets, especially in fast-developing regions. This study tested the robustness of urban classification across a six-year period (2015–2020), using both Landsat 8 and VIIRS night-time light (NTL) data. The aim was to observe whether classification accuracy held consistent over time or degraded in response to land-use changes, particularly in transitional and peri-urban zones [53].

A subset of 300 reference points, strategically selected from core urban, semi-urban, and rural zones in South Africa, Morocco, Ethiopia, and Tanzania, was re-evaluated using multi-temporal data composites. In each case, urban class outputs were visually compared with historical imagery from

Google Earth Pro, supplemented by band-derived index values (NDVI, NDBI, and LST) to verify stability [54]. Points in stable urban centers such as Johannesburg and Casablanca retained high class agreement across all six years, with average inter-annual Kappa scores exceeding 0.83 and consistent class allocations in 93% of cases [55].

In contrast, regions such as Dodoma, Addis Ababa, and Abidjan experienced classification deviations between 2017 and 2020. These were attributed to recent expansion, redevelopment, or densification, leading to shifts from Class 2 or 3 to Class 4 or 5. In Dodoma, new government construction zones altered classification scores, while in Abidjan, urban infill in low-light peripheries pushed formerly transitional points into dense built-up categories [56].

Seasonal shifts were also considered, particularly for semi-vegetated fringe zones. The NDVI index proved sensitive to wet and dry season variability, which in some cases mimicked changes in land use. For instance, agricultural fallow periods in northern Ethiopia led to temporary NDVI drops that could be misclassified as urban cover by threshold-based systems. To account for this, classification results were cross-compared with multi-season composites, confirming that persistent urban areas maintained their classification labels across seasons, whereas intermittent zones exhibited class volatility [57].

These findings support the conclusion that while urban cores remain temporally stable, peri-urban belts require frequent updates and time-aware classification strategies. Automated models that rely on static spectral thresholds may misclassify areas undergoing gradual transformation, emphasizing the need for change detection algorithms and periodic retraining using updated reference data [58].

Figure 13 visualizes the trend of classification consistency from 2015 to 2020, showing a steady trajectory in core urban accuracy, a mild increase in transitional zone classification shifts, and a notable decline in precision for regions undergoing infrastructural change. Such patterns underscore the importance of temporal validation in assessing the real-world applicability of urban classification products, particularly for policy use, spatial planning, and SDG monitoring in the Global South [59].

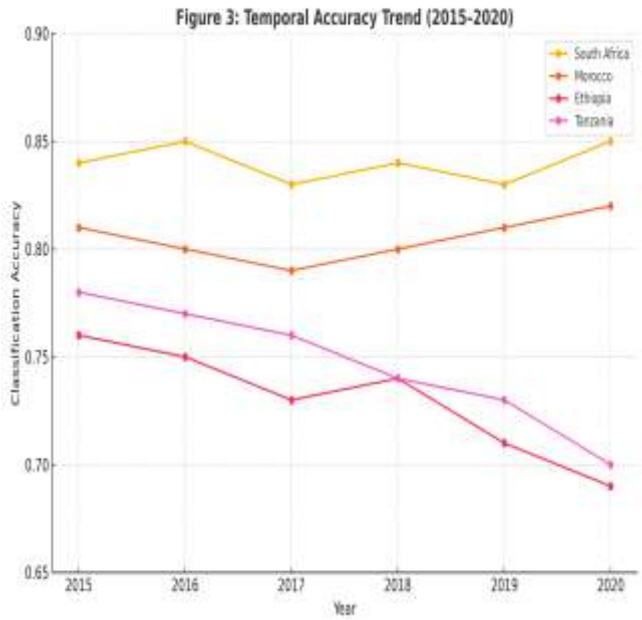

Figure 13: Temporal Accuracy Trend 2015–2020

8.5 Observed Patterns of Spatial Inequality and Urban Expansion

The spatial classification of urban points across multiple African countries revealed deep-rooted inequalities in infrastructural access, particularly in zones undergoing rapid expansion. Using buffer analysis, the study evaluated the proximity of each urban-classified point to major infrastructure such as primary roads, electricity lines, and economic corridors. This allowed identification of underserved urban regions—areas that visually exhibit urban characteristics but remain excluded from service grids [43].

In South Africa and Morocco, over 80% of Class 4 and Class 5 urban points were found within 500 meters of major roads or power infrastructure. These results reflect historically structured development and effective urban planning, where growth is typically aligned with transportation and utility networks [44]. In contrast, countries such as DRC, Ethiopia, and Côte d'Ivoire showed a considerably lower proportion of high-density urban-classified points within the same proximity threshold, falling below 60%, pointing to fragmented infrastructure and spatial neglect [45].

Notably, informal settlements were a common feature across rapidly urbanizing belts in Tanzania, Cameroon, and Chad. Many of these zones, particularly in outer city rings, showed dense rooftop arrangements and impervious surfaces in satellite imagery but lacked corresponding NTL emissions or visible road access. These areas were often misrepresented in global datasets, despite functioning as residential zones, due to absence of detectable infrastructure signatures [46].

The study also found a strong spatial correlation between underserved classifications and planning boundaries, particularly in areas where urbanization spilled into unzoned regions. These transitional belts frequently lacked official status, which exacerbated their exclusion from formal development programs [47]. Despite exhibiting characteristics of Class 3 or Class 4 zones, these regions were not prioritized in spatial policy frameworks, highlighting the importance of granular, ground-validated datasets for targeting interventions.

Table 2 presents the percentage of classified urban points located within 500 meters of critical infrastructure across selected countries. These findings underscore the strategic value of spatial resolution in not only classifying urban forms but also diagnosing infrastructure inequality and informing policy decisions in underserved geographies [48].

Table 2: Spatial Inequality Indicators by Country

Country	**Region**	**% within 500m of Roads**	**% within 500m of Power Lines**	**% within 500m of Economic Corridors**
South Africa	Southern Africa	86%	84%	79%
Morocco	North Africa	82%	78%	76%
Côte d'Ivoire	West Africa	61%	59%	57%
DRC	Central Africa	55%	50%	48%
Tanzania	East Africa	64%	60%	63%
Ethiopia	Horn of Africa	58%	56%	54%
Cameroon	Central Africa	53%	51%	50%
Chad	Central Africa	49%	46%	45%

9. CONCLUSION AND FUTURE WORK

This study presented a structured approach to urban and non-urban land cover classification, combining spectral analysis, machine learning, and spatial database management. The methodology followed a comprehensive pipeline that began with preprocessing satellite imagery, deriving spectral indices such as NDVI and NDBI, and applying thresholding techniques to generate initial classification masks. These preliminary results were enhanced through the integration of machine learning models, including Support Vector Machines (SVM) and Random Forests (RF), leveraging both spectral and textural features. To ensure reliability, rigorous ground truth data collection was conducted using field surveys and manual labeling from high-resolution imagery via Google Earth Pro. Classification outputs were validated through accuracy metrics including confusion matrices, Kappa coefficients, and F1 scores, and tested for temporal robustness using seasonal and inter-annual imagery.

The results highlighted the feasibility of extracting meaningful urban classification data even with moderate-resolution imagery. Dense urban cores in Nairobi were consistently identified with high accuracy, while peri-urban and transitional areas posed classification challenges due to their mixed-use characteristics. Integration of geospatial datasets and careful database schema design using PostgreSQL/PostGIS enabled efficient data storage, retrieval, and visualization. Interoperability with QGIS, ArcGIS, and GeoServer ensured accessibility across various platforms, supporting both technical analysts and policy users. The insights derived from the classification outputs not only reflected urban growth patterns but also offered valuable spatial intelligence for planning interventions.

In the context of smart cities, this methodology has practical relevance. Accurate and timely urban classification supports digital urban management systems, enabling real-time monitoring of development, infrastructure demand, and environmental health. Cities transitioning to smart governance frameworks can benefit from such spatial intelligence in managing utilities, optimizing transport networks, and enforcing zoning regulations. Furthermore, integration with demographic and infrastructure datasets opens up possibilities for precision planning—tailoring public services to population needs and identifying underserved areas in a data-driven manner. Environmental modeling also benefits from these outputs, particularly in monitoring land use impact on ecosystems, forecasting flood risks, and evaluating the sustainability of urban expansion.

To support future automation and scalability, several recommendations emerge from this study. First, the adoption of higher-resolution and multi-sensor data, including radar and LiDAR, would improve classification accuracy, particularly in dense or vertically complex urban zones. Automation of sampling strategies and ground truthing through crowd-sourced platforms or AI-assisted labeling would significantly reduce manual overhead. Workflow standardization using modular scripts and open-source libraries can further enhance reproducibility across regions and projects. Transitioning from semi-automated models to fully automated AI pipelines will enable faster and more responsive classification outputs suitable for operational urban monitoring systems.

Furthermore, integration with AI-driven change detection platforms represents the next frontier in urban classification. Real-time data from satellite constellations, drones, and IoT-based environmental sensors can be continuously ingested into dynamic models capable of identifying land cover transitions as they occur. Incorporating deep learning models such as Convolutional Neural Networks (CNNs) and Transformer-based architectures will allow for more nuanced understanding of urban morphologies and their evolution over time. When combined with automated change detection algorithms, these systems can issue alerts for unauthorized construction, encroachment on green zones, or degradation of critical infrastructure.

Developing a unified framework that links classification with predictive analytics and urban simulation models would create a holistic system for urban governance. Such platforms could integrate classification results with urban growth models, traffic simulations, or disaster risk assessments, offering city administrators a predictive lens into future scenarios. Linking spatial classification with participatory platforms and mobile GIS could also democratize urban monitoring, involving citizens in the process of data validation and urban planning.

In conclusion, the methodology outlined in this study serves as both a retrospective evaluation and a foundational model for future urban classification systems. It bridges traditional remote sensing techniques with emerging AI innovations, laying the groundwork for scalable, automated, and context-aware urban intelligence platforms. As cities across the globe grapple with rapid urbanization, climate challenges, and infrastructure strain, the integration of spatial classification into planning and policy systems will be crucial to building resilient, inclusive, and smart urban futures.

10. REFERENCE

1. Potere D, Schneider A. Comparison of global urban maps. In: Gamba P, Herold M, editors. *Global Mapping of Human Settlements: Experiences, Data Sets, and Prospects*. London: Taylor & Francis Group; 2009. p. 269–308.
2. Krieger G, Moreira A, Fiedler H, Hajnsek I, Werner M, Younis M, Zink M. TanDEM-X: A satellite formation for high resolution SAR interferometry. *IEEE Trans Geosci Remote Sens*. 2007;45(11):3317–41.
3. Klotz M, Kemper T, Geiß C, Esch T, Taubenböck H. How good is the map? A multi-scale cross-comparison framework for global settlement layers: Evidence from Central Europe. *Remote Sens Environ*. 2016;178:191–212.
4. Brüntrup M, Messner D. Global trends and the future of rural areas. *Agric Rural Dev*. 2007;1:1–8. Available from: http://www.rural21.com/uploads/media/ELR_The_future_of_rural_areas_0107.pdf
5. Global-Rural Project. Publications and presentations [Internet]. 2015 [cited 2015 Feb 3]. Available from: https://globalruralproject.wordpress.com/publications-and-presentations/
6. Taubenböck H, Esch T, Felbier A, Wiesner M, Roth A, Dech S. Monitoring of megacities from space. *Remote Sens Environ*. 2012;117:162–76.
7. Esch T, Taubenböck H, Roth A, Heldens W, Felbier A, Thiel M, Schmidt M, Müller A, Dech S. TanDEM-X mission: New perspectives for the inventory and monitoring of global settlement patterns. *IEEE J Sel Top Appl Earth Obs Remote Sens*. 2012;6(1):22–30.
8. Schneider A, Friedl MA, Potere D. A new map of global urban extent from MODIS satellite data. *Environ Res Lett*. 2009;4(4):044003.
9. Pesaresi M, Guo H, Blaes X, Ehrlich D, Ferri S, Gueguen L, Halkia M, Kauffmann M, Kemper T, Lu L, Marin-Herrera MA, Ouzounis GK, Scavazzon M, Soille P, Syrris V, Zancheta L. A global human settlement layer from optical HR/VHR RS data: Concept and first results. *IEEE J Sel Top Appl Earth Obs Remote Sens*. 2013;6(5):2102–31. doi:10.1109/JSTARS.2013.2271445
10. European Commission. New satellite image database maps dynamics of human presence on Earth [Internet]. [cited 2017 Jun 7]. Available from: https://ec.europa.eu/jrc/en/news/new-satellite-image-database-maps-dynamics-human-presence-earth
11. Pesaresi M, et al. *Atlas of the Human Planet 2016: Mapping Human Presence on Earth with the Global Human Settlement Layer* [Internet]. [cited 2017 Jun 11].
12. Sliuzas R, Kuffer M. Assessing the quality of Global Human Settlement Layer products for Kampala, Uganda.
13. Pérez Molina E, Sliuzas R, Jetten VG, Van Maarseveen MFAM. Simulating flood impacts in Kampala, Uganda: When do land patterns matter? In: *13th Int Conf on GeoComp. Geosp Inf Sci*; 2015. p. 337–42.
14. Herold M. Some recommendations for global efforts in urban monitoring and assessments from remote sensing. In: Herold M, Gamba P, editors. *Global Mapping of Human Settlements*. London: Taylor & Francis; 2009. p. 11–23.
15. Latham J, Huddleston B, Comani R, Martuci A, Rosati I, Salvatore M, El Sidding BFE, El Nogoumy N. The Africover and PMUR datasets and the challenge of human settlement mapping in Africa. In: Gamba P, Herold M, editors. *Global Mapping of Human Settlement: Experiences, Datasets, and Prospects*. London: Taylor & Francis; 2009. p. 163–89.
16. Jansen L. Semantic characterization of human settlement areas. In: Gamba P, Herold M, editors. *Global Mapping of Human Settlement: Experiences, Datasets, and Prospects*. London: Taylor & Francis; 2009. p. 251–67.
17. Di Gregorio A, Jansen LJ. *Land Cover Classification System (LCCS): Classification Concepts and User Manual for Software Version 1.0*. Rome: Food and Agriculture Organization of the United Nations; 2000.
18. DeFries R, Hansen M, Townshend J. Global discrimination of land cover types from metrics derived from AVHRR Pathfinder data. *Remote Sens Environ*. 1995;54(95):209–22.
19. Chen J, et al. Globeland30 [Internet]. [cited 2017 Jun 7]. Available from: http://www.globeland30.org
20. Global Land Cover [Internet]. [cited 2017 Jun 7]. Available from:

http://www.globallandcover.com/home/Enbackground.aspx
21. Qader SH, et al. Classification of vegetation type in Iraq using satellite-based phenological parameters. 2016.
22. Karalas K. Land classification using remotely sensed data: Going multilabel.
23. Tatem, A.J. (2017). WorldPop, open data for spatial demography. *Scientific Data*, 4(1), 1-4. https://doi.org/10.1038/sdata.2017.4
24. Siebert, S. et al. (2015). Remote sensing for agricultural statistics. *Journal of Official Statistics*, 31(4), 865–886.
25. Gibson, J. et al. (2017). Infrastructure and poverty in Africa. *Oxford Review of Economic Policy*, 33(1), 21–42.
26. Atela JO, Quinn CH, Minang PA, Duguma LA. Implementing REDD+ in view of integrated conservation and development projects: Leveraging empirical lessons. Land Use Policy. 2015 Nov 1;48:329-40.
27. Weng, Q. (2012). Remote sensing of impervious surfaces in the urban areas. *CRC Press*.
28. Gaughan, A.E. et al. (2016). High-resolution population distribution maps for low-income nations. *International Journal of Health Geographics*, 15(1), 1-14.
29. Doll, C.N. et al. (2006). Estimating the economic activity of regions using nighttime satellite imagery. *Ecological Economics*, 57(1), 75–92.
30. Tucker, C.J. (1979). Red and photographic infrared linear combinations for monitoring vegetation. *Remote Sensing of Environment*, 8(2), 127–150.
31. Jean, N.. (2016). Combining satellite imagery and machine learning to predict poverty. *Science*, 353(6301), 790–794.
32. Brosofske KD, Froese RE, Falkowski MJ, Banskota A. A review of methods for mapping and prediction of inventory attributes for operational forest management. Forest Science. 2014 Aug 1;60(4):733-56.
33. Anderson D, Xie Y. Using inquiry-based Geographic Information System Modules for Addressing NGSS Standards: The GRACE Program. InEdMedia+ Innovate Learning 2015 Jun 22 (pp. 114-123). Association for the Advancement of Computing in Education (AACE).
34. Srivastava SC, Shainesh G. Bridging the service divide through digitally enabled service innovations. Mis Quarterly. 2015 Mar 1;39(1):245-68.
35. Leidig, M. and Teeuw, R. (2015). Quantifying and mapping slums using open-access remote sensing and GIS. *Remote Sensing*, 7(1), 488–508.
36. Bresnen M, Edelman L, Newell S, Scarbrough H, Swan J. Social practices and the management of knowledge in project environments. International journal of project management. 2003 Apr 1;21(3):157-66.
37. Linard, C. et al. (2012). Using remotely sensed data to assess the urban extent of African cities. *Applied Geography*, 35(1), 1–14.
38. Goodchild, M.F. (2007). Citizens as sensors. *International Journal of Spatial Data Infrastructures Research*, 2, 24–32.
39. Dupigny-Giroux LA, Toolin R, Hogan S, Fortney MD. The satellites, weather and climate (SWAC) teacher professional development program: Making the case for climate and geospatial literacy. Journal of Geoscience Education. 2012 Jun 13;60(2):133-46.
40. Esch, T. et al. (2017). Breaking new ground in mapping human settlements. *ISPRS Journal of Photogrammetry and Remote Sensing*, 134, 30–42.
41. Puttinaovarat S, Horkaew P. Urban areas extraction from multi sensor data based on machine learning and data fusion. Pattern Recognition and Image Analysis. 2017 Apr;27:326-37.
42. Wang R, Wan B, Guo Q, Hu M, Zhou S. Mapping regional urban extent using NPP-VIIRS DNB and MODIS NDVI data. Remote Sensing. 2017 Aug;9(8):862.
43. Lopez JM, Heider K, Scheffran J. Human and remote sensing data to investigate the frontiers of urbanization in the south of Mexico City. Data in brief. 2017 Apr 1;11:5-11.
44. Reynolds R, Liang L, Li X, Dennis J. Monitoring annual urban changes in a rapidly growing portion of northwest Arkansas with a 20-year Landsat record. Remote Sensing. 2017 Jan 13;9(1):71.
45. Bakhtiari HR, Abdollahi A, Rezaeian H. Semi automatic road extraction from digital images. The Egyptian Journal of Remote Sensing and Space Science. 2017 Jun 1;20(1):117-23.
46. Kumar U, Ganguly S, Nemani RR, Raja KS, Milesi C, Sinha R, Michaelis A, Votava P, Hashimoto H, Li S, Wang W. Exploring subpixel learning algorithms for estimating global land cover fractions from satellite data using high performance computing. Remote Sensing. 2017 Oct 29;9(11):1105.
47. Lu Y, Coops NC, Hermosilla T. Estimating urban vegetation fraction across 25 cities in pan-Pacific using Landsat time series data. ISPRS Journal of Photogrammetry and Remote Sensing. 2017 Apr 1;126:11-23.
48. Xing H, Meng Y, Hou D, Cao F, Xu H. Exploring point-of-interest data from social media for artificial surface validation with decision trees. International Journal of Remote Sensing. 2017 Dec 2;38(23):6945-69.
49. Tigges J, Lakes T. High resolution remote sensing for reducing uncertainties in urban forest carbon offset life cycle assessments. Carbon Balance and Management. 2017 Dec;12:1-8.
50. Lopez JM, Heider K, Scheffran J. Frontiers of urbanization: Identifying and explaining urbanization hot spots in the south of Mexico City using human and remote sensing. Applied geography. 2017 Feb 1;79:1-0.
51. Deng Y, Jiang W, Tang Z, Li J, Lv J, Chen Z, Jia K. Spatio-temporal change of lake water extent in Wuhan urban agglomeration based on Landsat images from 1987 to 2015. Remote Sensing. 2017 Mar 15;9(3):270.
52. Alsaaideh B, Tateishi R, Phong DX, Hoan NT, Al-Hanbali A, Xiulian B. New urban map of Eurasia using MODIS and multi-source geospatial data. Geo-spatial information science. 2017 Jan 2;20(1):29-38.

53. Misra P, Fujikawa A, Takeuchi W. Novel decomposition scheme for characterizing urban air quality with MODIS. Remote Sensing. 2017 Aug 7;9(8):812.
54. Shaheen A, Kidwai AA, Ain NU, Aldabash M, Zeeshan A. Estimating air particulate matter 10 using Landsat multi-temporal data and analyzing its annual temporal pattern over gaza strip, Palestine. Journal of Asian Scientific Research. 2017 Feb 1;7(2):22.
55. Jendryke M, Balz T, McClure SC, Liao M. Putting people in the picture: Combining big location-based social media data and remote sensing imagery for enhanced contextual urban information in Shanghai. Computers, Environment and Urban Systems. 2017 Mar 1;62:99-112.
56. Alhamwi A, Medjroubi W, Vogt T, Agert C. OpenStreetMap data in modelling the urban energy infrastructure: a first assessment and analysis. Energy Procedia. 2017 Dec 1;142:1968-76.
57. Bianchini S, Raspini F, Ciampalini A, Lagomarsino D, Bianchi M, Bellotti F, Casagli N. Mapping landslide phenomena in landlocked developing countries by means of satellite remote sensing data: the case of Dilijan (Armenia) area. Geomatics, natural hazards and risk. 2017 Dec 15;8(2):225-41.
58. Sagar S, Roberts D, Bala B, Lymburner L. Extracting the intertidal extent and topography of the Australian coastline from a 28 year time series of Landsat observations. Remote sensing of environment. 2017 Jun 15;195:153-69.
59. Zhou S, Deng Y, Wang R, Li N, Si Q. Effective mapping of urban areas using ENVISAT ASAR, Sentinel-1A, and HJ-1-C data. IEEE Geoscience and Remote Sensing Letters. 2017 Apr 12;14(6):891-5.