

Geospatial Mapping of Community Level Electrification in Nigeria Using VIIRS Nighttime Light Data with Threshold Optimization and Ground Validation



Y. A. Badamasi^{1, 2}, S. Adeshina^{1*}, A. D. Usman²

¹Department of Electrical & Electronics Engineering, Nile University, Abuja, Nigeria.

² Department of Electrical and Computer Engineering, Faculty of Engineering, Baze University, Abuja, Nigeria.



ABSTRACT: Reliable and spatially detailed information on electricity access is essential for effective rural electrification planning in Nigeria; however, existing statistics are typically reported at aggregated administrative levels, thereby masking important community-scale disparities. This study develops a high-resolution, cluster-level electrification mapping framework for Nigeria using satellite-derived night-time light observations, integrated with a statistically optimized radiance threshold and field-based validation. The 2024 Visible Infrared Imaging Radiometer Suite (VIIRS) Day/Night Band (DNB) radiance data were processed and corrected for gas flaring using VIIRS Nightfire products, and were subsequently applied to 22,696 geospatially defined community clusters in Nigeria. Candidate radiance thresholds within the literature-supported range of $0.25\text{--}0.50 \text{ nW}\cdot\text{cm}^{-2}\cdot\text{sr}^{-1}$ were systematically evaluated using a grid-search approach that maximized the F1-score, and electrification status was assigned through cluster-level sampling at the resulting optimal threshold. The classification was validated against a combined reference dataset comprising Rural Electrification Agency mini-grid records, field-verified unelectrified communities, and confirmed electrified sites, using confusion matrix-based performance measures. Classification performance was most balanced at a threshold of $0.35 \text{ nW}\cdot\text{cm}^{-2}\cdot\text{sr}^{-1}$, giving a precision of 0.69, a recall of 0.86, and an F1-score of 0.77, with 79.45 percent of clusters classified as unelectrified. The study contributes an optimized threshold-selection framework, a Nigeria-specific gas-flare correction approach, and a high-resolution electrification dataset, supporting electrification planning and policy formulation.

KEYWORDS: Electrification planning, VIIRS Night-time Light, VIIRS Night Fire, Geospatial analysis, Energy access, Sub-Saharan Africa.

[Received Mar. 24, 2026; Revised Apr. 04, 2026; Accepted Apr. 07, 2026]

Print ISSN: 0189-9546 | Online ISSN: 2437-21107

I. INTRODUCTION

The availability of sustainable electricity is one of the principal developmental problems in Sub-Saharan Africa, in which a large percentage of the population still lacks access to modern energy services (IEA et al., 2024). Achieving Sustainable Development Goal 7 (SDG7), which targets universal access to affordable, reliable, and sustainable energy, requires accurate and spatially explicit information on electrification status. However, conventional survey-based approaches are limited by high cost, infrequent updates, and lack of geographic detail, making them inadequate for large-scale and dynamic electrification planning (IEA, 2023).

Night-time light (NTL) remote sensing has now become an established proxy of human activity and infrastructure development, and the use of the technique has been demonstrated in urbanization mapping, economic estimation, and energy consumption analysis. Numerous studies have demonstrated that electrified settlements exhibit higher radiance values than unelectrified ones, enabling threshold-based classification of electrification status from satellite observations (Doll and Pachauri, 2010; Min et al., 2013). Building on this foundation, (Falchetta et al., 2019) developed high-resolution electrification datasets for Sub-Saharan Africa, while (Giacomo Falchetta et al., 2020) analyzed spatial

inequalities in electrification access. Similarly, (Khavari et al., 2021) integrated population clusters with geospatial electrification modelling to improve planning relevance, and (Gao et al., 2022) proposed a global framework for identifying electrified and unelectrified built-up areas using Visible Infrared Imaging Radiometer Suite (VIIRS) derived thresholds, demonstrating high classification accuracy.

However, despite these advances, existing approaches remain insufficient for localized electrification planning. To begin with, the majority of research works assume globally or regionally specified radiance thresholds, which are not sufficient to reflect the nonhomogeneous radiance features of rural settlements, especially in countries like Nigeria, where the level of electricity consumption is low and intermittent. Consequently, small power yet truly electrified communities are often mistaken to be unelectrified. For example, while (Falchetta et al., 2019) and (Gao et al., 2022) provide valuable large-scale electrification maps, their reliance on generalized thresholding frameworks limits their sensitivity to weak radiance signals typical of rural environments. Similarly, (Khavari et al., 2021) improved spatial representation through clustering, but their framework primarily focuses on spatial demand representation rather than explicitly addressing the radiometric uncertainties associated with night-time light data

*Corresponding author: steve.adeshina@nileuniversity.edu.ng

doi: <http://doi.org/10.63746/njtd.v23i3.4579>

and does not incorporate statistically validated threshold selection for electrification classification.

A review of existing studies further indicates a convergence toward VIIRS radiance thresholds in the range of $0.25\text{-}0.50 \text{ nW}\cdot\text{cm}^{-2}\cdot\text{sr}^{-1}$ for distinguishing unelectrified from electrified areas (Coesfeld et al., 2020; Elvidge et al., 2021; Falchetta et al., 2019; Yuan et al., 2022, 2019). These levels are usually based on empirical measurements of sensor noise floors and radiance distributions in non-electrified areas and usually checked against survey data and population samples. Nonetheless, the choice of threshold is still not uniform across the literature as various researchers use different cut-offs in the range depending on their methodology and study situation. Worse still, the current methods lack a statistically based system to determine an ideal threshold, rather they use either heuristic or transferred values, which might not be applicable across regions.

In addition, the reliability of NTL data are subject to several sources of uncertainty. VIIRS sensors retrieve radiance, which is influenced by atmospheric conditions, lunar illumination, sensor geometry, vegetation cover and cloud contamination (Wang et al., 2021) making the data subject to variability. All these ambiguities are more specific to the conditions of low radiance, where the line between electrified and unelectrified space is less distinct. While filtering techniques such as the pseudo-invariant feature method (PFM) proposed by (Yuan et al., 2019) improves the detection of built-up areas, they do not directly address electrification classification and remain sensitive to threshold selection under noisy conditions. In addition, other sources of complexity are industrial and environmental light sources. In regions such as Nigeria, gas flaring associated with oil production, bush burning and other combustion-related activities produces persistent high-intensity radiance that can be misinterpreted as electrification, leading to significant false-positive classifications. Despite its importance, this source of error is rarely explicitly addressed in existing electrification mapping frameworks, thereby limiting their applicability in gas-flaring-prone regions.

Moreover, many remote sensing approaches focus on identifying settlements rather than explicitly classifying electrification status. Multispectral classification methods which are deployed to satellite images including Sentinel-2 have been effectively utilized in outlining built up districts and rural settlements (Adelabu and Olusola, 2021). However, these approaches do not directly capture electricity access and therefore require integration with additional datasets for meaningful electrification assessment. Consequently, a critical gap remains in developing a validated, community-scale electrification classification framework that integrates high-resolution settlement units, statistically robust threshold selection, and context-specific noise mitigation.

In order to overcome these constraints, this paper uses a cluster-based representation of rural settlements based on the geospatial model of (Husein et al., 2024), which consists of 22,696 settlement clusters in Nigeria. Building on this foundation, the study integrates VIIRS night-time light data with a systematic threshold optimization approach based on F1-score maximization and validates the resulting classification using a combined dataset of Rural Electrification Agency (REA) installed mini-grids records and field observations. This study has three-fold contributions. First, it provides a granular, community-level electrification mapping

framework based on 22,696 settlement clusters, enabling improved spatial resolution for planning applications. Second, it introduces a Nigeria-specific calibration approach by incorporating VIIRS Night-fire data to identify and remove gas flaring and other combustion-related noise sources, thereby reducing false-positive electrification signals. Third, it creates a statistically robust framework of threshold optimization, by maximizing F1-score, and testing on ground-truth data. Together, these contributions provide a reproducible and planning-relevant methodology for high-resolution electrification mapping in Nigeria.

II. THEORETICAL ANALYSIS

A. Night-Time Radiance Detection

Satellite nighttime light radiance serves as a robust proxy for assessing electrification at the community level, based on the premise that artificial light intensity emitted from electrified settlements shows measurable radiance emissions detectable by spaceborne sensors. The VIIRS Day/Night Band (DNB) can be regarded as especially useful because it is more sensitive to the low-light conditions, which allows observing anthropogenic light intensity patterns linked to electricity access. This capability makes VIIRS data especially suitable for electrification studies in regions where ground truth data are limited or unavailable, especially in low-light areas like Sub-Saharan Africa, which are mostly developing countries. The Mathematical relationship governing the observed radiance is expressed in equation (1)

$$L(x, y) = f(I_{light}, A, d, \theta) \quad (1)$$

The observed radiance $L(x, y)$ is dependent on the light intensity I_{light} , illuminated surface area A , θ and the distance between the light source and the satellite sensor d . Equation (1) is given in a generalized functional form instead of defining a deterministic physical model. The equation in the framework of the current research is conceptual in that it demonstrates that the artificial light intensity is only one of the factors that affect the satellite-observed radiance, and other elements of the environment and sensors also play a role. Since these interactions are complex and variable under the conditions of nature and in low-radiance rural environments, in particular, it is not feasible to explicitly parameterize every component. Rather, the empirical methodology of the study is that of empirical classification, whereby the overall effect of the aforementioned factors is implicitly represented by measured radiance values and measured using threshold-based techniques. This justifies the use of simplified functional representation rather than an explicit formulation of the analysis.

The observed radiance at a given pixel location is denoted as $L(x, y)$, and is influenced by several interacting factors. These are the intensity of artificial light I_{light} , light surface area A and distance between the light source and the satellite sensor d . Moreover, atmospheric conditions and sensor viewing geometry, denoted by θ , influences the attenuation and scattering of emitted light before it is captured. Formulation emphasizes that the radiance captured is not only a variable of the presence of electricity, but also the environmental and geometrical factors that can affect the signal strength. Nevertheless, the intensity of artificial light remains strongly correlated with electricity access making $L(x, y)$ a meaningful and widely adopted indicator for

mapping electrification patterns (Elvidge et al., 2017; G Falchetta et al., 2020; Li et al., 2020). This relationship underpins the use of nighttime light data in large scale electrification assessments.

B. Settlement Radiance Aggregation

There are various sources of artificial light intensity in a settlement, and they include residential lighting, street lighting, commercial lighting and community infrastructure. Consequently, the brightness measured by the satellite sensor is the aggregate of these various sources of lights in a particular pixel footprint. It is particularly notable in the case of heterogeneous settlements, in which the intensity and distribution of lighting are spatially heterogeneous. This relationship can be approximated using a radiance aggregation model, as shown in equation (2).

$$L_p = \sum_{i=1}^n I_i \cdot \eta_i \quad (2)$$

The observed pixel radiance L_p can be described as the sum of contributions from multiple light sources, each weighted by atmospheric transmission and sensor response factors η_i that take into account signal attenuation and detection efficiency, the summation of all contributing sources n . This expression is based on the reality that the radiance of a satellite is not only driven by the existence of lighting, but also sensing and environmental conditions. As a result, the electrified settlements are expected to have a high radiance value because of many and rather powerful light sources. Unlike this, unelectrified or low electrified regions generate weak radiance signals that can often be close to background noise levels. This distinction forms the basis for using radiance threshold to differentiate between electrified and unelectrified settlements particularly in large scale spatial analysis or electrification planning.

C. Radiance Threshold Classification

Electrification detection using satellite-derived night-time light data commonly relies on threshold-based classification methods. In this approach, the observed radiance $L(x, y)$ at a given spatial location is compared against a predefined threshold to infer electrification status. Settlements with radiance values exceeding the threshold are classified as electrified, while those below the threshold are considered unelectrified. This binary representation provides a simple yet effective means of translating continuous radiance values into discrete electrification states, particularly for large-scale spatial analysis.

In this study, the observed radiance $L(x, y)$ is compared against a predefined scalar threshold T which has the same physical units as radiance ($\text{nW} \cdot \text{cm}^{-2} \cdot \text{sr}^{-1}$). The threshold T is constant across space and does not vary with location. This ensures dimensional consistency while maintaining a uniform classification criterion across all settlement clusters. The classification rule is expressed as:

$$E(x, y) = \begin{cases} 1 & \text{if } L(x, y) \geq T \\ 0 & \text{if } L(x, y) < T \end{cases} \quad (3)$$

where $E(x, y)$ in equation (3) represents electrification status at location (x, y) , $L(x, y)$ is the observed radiance and T is the radiance threshold. Although spatially varying thresholds

$L_{th}(x, y)$ may be adopted in more advanced classification frameworks to account for local variability, this study employs a uniform threshold to ensure methodological simplicity, consistency, and comparability across all regions. While this method is computationally efficient and widely adopted, its accuracy is highly sensitive to the choice of threshold. An improperly selected threshold may lead to significant misclassification, particularly in regions characterized by low or heterogeneous lighting intensity. This highlights the importance of data-driven threshold optimization, as implemented in this study.

D. Classification Accuracy Metrics

The selection of the threshold T is therefore critical and requires careful calibration to balance false positives (misclassifying non-electrified areas as electrified) and false negatives (failing to detect electrified areas). Achieving this balance is essential for ensuring reliable classification outcomes, especially in rural context where radio signals may be weak or inconsistent (Min et al., 2013; Jean et al., 2016; Falchetta et al., 2019). Confusion matrix measures are used to measure the performance of classification, and they are true positives (TP), true negatives (TN), false positives (FP), and false negatives (FN). These measures are the foundation of computing the standard measures of performance, which include: precision, recall, accuracy and the F1-score, as respectively

$$P = \frac{TP}{TP+FP} \quad (4)$$

$$R = \frac{TP}{TP+FN} \quad (5)$$

$$Acc = \frac{TP+TN}{TP+TN+FP+FN} \quad (6)$$

$$F1 = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (7)$$

where P , R , and Acc represent precision, recall, and accuracy, respectively. Among these metrics, the F1-score is particularly valuable in scenarios with imbalanced class distributions, as it provides balanced measures of precision and recall. This is especially relevant for electrification studies, where unelectrified settlements often dominate the dataset (Husein et al., 2024), potentially biasing conventional accuracy measures.

Equations (4)-(7) define the performance metrics used to evaluate classification accuracy and are directly applied in the methodology for threshold optimization. Specifically, these metrics are computed for each candidate radiance threshold using the validation dataset described in Section III.E. Among these, the F1-score plays a central role in this study, as it provides a balanced measure of precision and recall, which is particularly important in imbalanced classification scenarios where unelectrified settlements dominate. The optimization problem defined in equation (8) uses the F1-score as the objective function to determine the optimal radiance threshold. Therefore, these equations are not merely theoretical definitions but form the analytical foundation for evaluating classification performance and selecting the optimal threshold

applied in the results presented in Section IV. To determine the optimal classification threshold, the problem is formulated as an optimization task in which the threshold T is selected to maximize the F1-score across evaluated radiance values.

$$T^* = F1(T)_{max} \quad (8)$$

where T^* represents the optimal radiance threshold and $F1(T)$ is the F1-score corresponding to the threshold T . This optimization criterion ensures that the selected threshold achieves the best trade-off between precision and recall while minimizing classification errors. As a result, it provides a more robust and data-driven basis for electrification mapping compared to arbitrary or fixed threshold.

E. Cluster Representation of Settlements

To operationalize electrification mapping at the community scale, rural settlements are represented as cluster centroids derived from geospatial population datasets. This cluster-based representation enables electrification classification to be performed on discrete community units rather than aggregated administrative boundaries, thereby improving the spatial granularity and practical relevance of the analysis. Such an approach is particularly important in rural context, where administrative units may mask significant intra-regional variation in electrification status. The electrification status of cluster (c_i) is sampled directly from the radiance raster at the centroid coordinates (x_i, y_i), expressed in equation (9).

$$E(c_i) = E(x_i, y_i) \quad (9)$$

Where c_i denotes the centroid of cluster i , and $E(x_i, y_i)$ represent the electrification value obtained from the radiance raster at that location. This Formulation provides a direct linkage between remotely sensed radiance data and spatial

representation of human settlements. By adopting this approach, the analysis captures electrification patterns at a resolution that aligns more closely with real world communities. This is particularly advantageous for identifying localized pockets of unelectrified settlements that may be overlooked in coarser analysis.

More broadly, the integration of nighttime lights satellite radiance data with geospatial settlement cluster datasets forms a robust foundation for large scale electrification mapping. This synergy combines the strength of remote sensing in capturing dynamic light conditions with the spatial precision of population-based settlement modeling that are highly relevant for energy planning policy formulation and infrastructure development. In the context of Nigeria such a framework provides a scalable and data-driven approach for monitoring electrification progress and supporting targeted interventions (Falchetta et al., 2020; IEA, 2023).

III. MATERIALS AND METHODS

This study adopts a quantitative geospatial framework to categorise the electrification level of settlement clusters in Nigeria. The method combines both satellite derived night time lights NTL radiance data and high-resolution settlement cluster data sets that allow the measurement of electrification at a community level. This approach is particularly necessary in regions of data scarcity, where traditional ground-based electrification data are scarce or not available. To enhance the consistency of the classification, the noise to signal ratio that is typically typified with the low radiance rural backdrop is minimized by the optimised radiant threshold technique. This is to make sure that useful but insignificant light intensity indicators are not confused as unelectrified areas. A systematic view of the data preprocessing and analysis pipeline is illustrated in Figure 1.

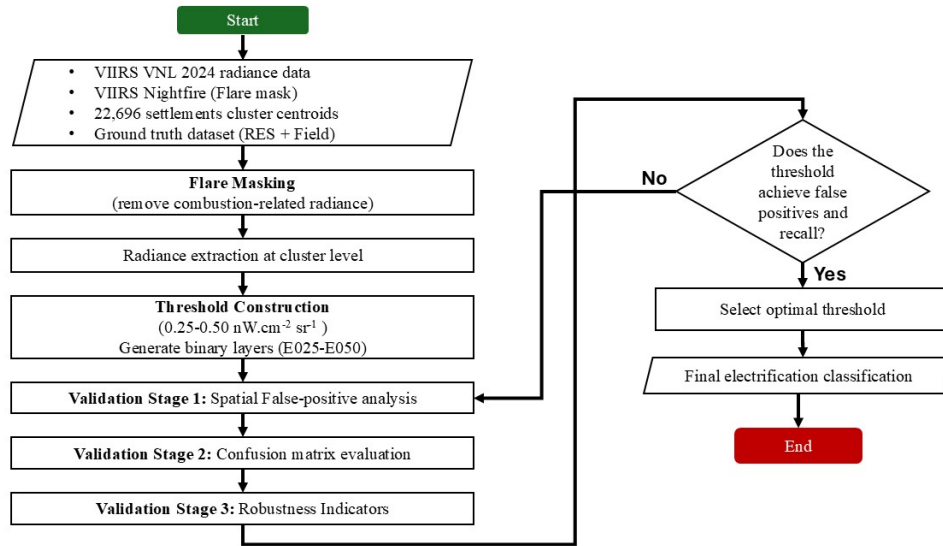


Figure 1 Flowchart of Electrification Classification Framework

Raster preprocessing, spatial masking, and radiance sampling operations were conducted in QGIS version (3.34.15), including the extraction of pixel values at settlement cluster locations. The VIIRS Nighttime Light (VNL) and

VIIRS Nightfire (VNF) products, produced by the Earth Observation Group (EOG) at the Colorado School of Mines (Earth Observation Group, 2024a, 2024b), served as the primary data sources. QGIS was then utilized for efficient

visualization, and large-scale processing of these datasets. This combination of tools and data sources ensured both computational efficiency and reproducibility in handling large national-scale geospatial datasets.

A. Study Area and Settlement Clusters

This study is conducted in Nigeria, which is situated in the western part of the African continent. It is marked by varying geographic and socio-economic factors, where a high number of its population can be found in the rural areas, which experience consistent electricity accessibility problems. Such differences render Nigeria as an appropriate case study for large-scale electrification analysis and upgrades. The unit of analysis comprises spatially defined settlement clusters, comprises settlement clusters that make up rural settlements in the entire nation. The data set includes 22,696 settlements

centroids (Husein et al., 2024), each settlement centroid is assigned to a unique community location. A “settlement cluster” refers to a spatial aggregation of built-up structures and population derived from high-resolution gridded datasets, representing a discrete rural community. The clusters used in this study were obtained from (Husein et al., 2024), where they were generated through the integration of population density and built-up area datasets at approximately 100 m resolution. These centroids allow the evaluation of identification status to be done at a high spatial resolution, thereby providing a more realistic and detailed electricity access compared to aggregated administrative units. Figure 2 illustrates the spatial distribution of settlement clusters classified as unelectrified, highlighting the scale and dispersion of the electrification deficit across the study area and providing spatial context for the cluster-level classification approach adopted in this study.

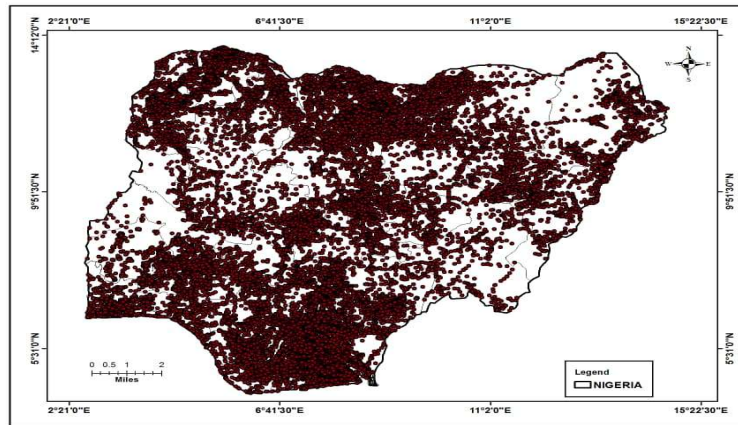


Figure 2 Unelectrified Clusters (Husein et al., 2024)

B. Input Data and VIIRS Processing

The analysis integrates multiple geospatial datasets to support electrification classification at the community level. A total of 22,696 settlement cluster centroids obtained from (Husein et al., 2024) were used as the spatial framework for radiance sampling and analysis. Night-time radiance data were obtained from the VIIRS Day/Night Band (DNB) Version 2.2 annual average composite for 2024 (Earth Observation Group, 2024a; Elvidge et al., 2021). This dataset was selected due to its improved radiometric calibration, reduced stray light contamination, and integration of observations from both the Suomi-NPP and NOAA-20 satellites, which enhance temporal consistency and data reliability. Radiance values are expressed in nanowatts per square centimeter per steradian ($nW \cdot cm^{-2} \cdot sr^{-1}$), enabling standardized quantitative comparison across regions.

To reduce contamination from non-anthropogenic light sources, VIIRS Nightfire (VNF) 2024 data (Earth Observation Group, 2024b) were incorporated to identify combustion-related emissions such as gas flaring and biomass burning. A temporally subsampled set of VNF observations was used to construct an annual gas flare mask, while the detailed justification for the sampling strategy is provided in Section III.C. The VIIRS Nightfire dataset provides geolocated point detections of active combustion sources. To convert these point detections into a spatial mask, each Nightfire point was buffered using a radius of approximately 500 m, corresponding to the spatial footprint of VIIRS DNB observations. The buffered points were merged to form a continuous polygon layer representing combustion-affected areas. This polygon layer was subsequently rasterized and applied as a mask to the VIIRS radiance dataset, such that all pixels intersecting the mask were excluded from further analysis. The input dataset is as shown in Table 1.

Table 1 Input Dataset

Dataset	Description	Source
Clusters	Settlement cluster centroids	Husein et al., 2024 Click or tap here to enter text.
VIIRS VNL 2024	Annual night-time radiance	Earth Observation Group, 2024a; Elvidge et al., 2021
VIIRS Nightfire 2024	Gas flare detection	Earth Observation Group, 2024b

C. Spatial Resolution, Sampling Strategy, and Computational Environment

All geospatial processing and analysis were conducted using QGIS (version 3.34.15). The excel file was uploaded as delimited text layer using the coordinates (longitude and latitude) and Raster preprocessing, spatial sampling, and map visualization were performed in QGIS. The analysis integrates datasets with differing spatial resolutions. The settlement cluster dataset was derived at an approximate spatial resolution of 100 m, representing the scale of rural settlement delineation, while the VIIRS Day/Night Band radiance data has a nominal spatial resolution of approximately 500 m at nadir. To reconcile these differences, radiance values were extracted at each settlement cluster centroid using nearest-pixel sampling. This approach preserves the discrete spatial representation of communities while ensuring that sampled radiance values correspond directly to observed satellite measurements without introducing interpolation or smoothing effects.

Due to the high temporal frequency and data volume of VIIRS Nightfire (VNF) observations, a systematic temporal subsampling strategy was adopted. Specifically, VNF data were extracted for the 1st, 10th, and 20th day of each month, resulting in a total of 36 observation days for the year 2024. These dates were selected to provide evenly distributed temporal coverage within each month, ensuring representation of intra-month variability in combustion activity while maintaining computational efficiency (Wang et al., 2021). This sampling approach is consistent with established super-resolution and filtering practices used to catalog persistent infrared heat sources while eliminating false positives caused by atmospheric glow (Zhizhin et al., 2026). By utilizing infrared detectors that respond exclusively to heat emissions, this method ensures that transient thermal anomalies are separated from the persistent industrial signals required for accurate electrification masking (World Bank, 2024). The objective of this sampling is not to detect all transient combustion events, but to identify persistent or recurrent sources of radiance that could bias electrification classification. Gas flaring, which represents the primary source of non-residential radiance contamination in Nigeria, is typically continuous or frequently recurring rather than sporadic. Consequently, repeated detection across multiple temporally spaced observations is sufficient to identify stable combustion locations. The adequacy of the 36-day sampling framework was further supported by sensitivity testing using higher-frequency subsets of VNF data, which showed no significant difference in the identification of persistent combustion sources. This confirms that the adopted sampling strategy provides a reliable and computationally efficient approximation of combustion-related radiance patterns relevant to electrification mapping.

Radiance values were extracted from the VIIRS DNB raster using point-based sampling at the centroid of each settlement cluster. Specifically, the radiance value corresponding to the nearest raster pixel center was assigned to each cluster. No interpolation, resampling, or spatial averaging was applied during this process. This nearest-pixel sampling approach ensures that extracted values directly represent observed satellite radiance measurements, avoiding the introduction of artificial smoothing or estimation bias that may arise from interpolation or buffer-based averaging

methods. Given the coarse spatial resolution of VIIRS data 500 m, this approach provides a consistent and reproducible representation of radiance at the community level.

Centroid-based sampling was adopted in preference to zonal statistics or buffer-based averaging for both conceptual and methodological reasons. First, the settlement clusters used in this study represent discrete community units, and centroid sampling preserves this representation by assigning a single radiance value to each cluster, consistent with the objective of community-level electrification classification. Second, the spatial resolution of VIIRS DNB data 500 m is relatively coarse compared to the spatial extent of many rural settlements. Applying zonal statistics or buffer averaging in such contexts may introduce spatial smoothing effects, blending radiance signals from adjacent areas, including non-settlement regions or nearby light sources. This can lead to misrepresentation of the actual lighting conditions within a settlement, particularly in heterogeneous rural environments. Third, centroid sampling avoids the introduction of aggregation bias and ensures that each cluster is evaluated independently based on observed radiance values. This is particularly important for threshold-based classification, where small variations in radiance values can influence electrification status. Finally, centroid-based extraction offers significant computational efficiency when processing large national-scale datasets. Given the large number of settlement clusters (22,696), this approach enables scalable and reproducible analysis without compromising classification accuracy.

D. Radiance Threshold Evaluation and Radiance Sampling

Electrification classification was implemented using a radiance threshold approach, as introduced in the theoretical framework (Section II). The candidate threshold range ($0.25\text{--}0.50 \text{ nW}\cdot\text{cm}^{-2}\cdot\text{sr}^{-1}$) was derived from values consistently reported in the literature for distinguishing electrified and unelectrified areas. Within this range, a grid-search-based evaluation was conducted at 0.05 increments ($0.25, 0.30, 0.35, 0.40, 0.45, \text{ and } 0.50 \text{ nW}\cdot\text{cm}^{-2}\cdot\text{sr}^{-1}$) to systematically assess classification performance across thresholds. This approach avoids reliance on arbitrary threshold selection and instead enables data-driven identification of the optimal threshold. For each candidate threshold, binary radiance surfaces were generated by classifying pixels with radiance values above the threshold as electrified and those below as unelectrified. This multi-threshold evaluation allows detailed analysis of classification sensitivity, particularly in low-radiance rural environments where signal ambiguity is common.

The electrification status of each settlement cluster was then determined by sampling the thresholded raster at the centroid location of each cluster, consistent with the cluster representation model described in Section II. This sampling approach establishes a direct link between satellite-derived radiance measurements and discrete community locations, enabling accurate assignment of electrification status at the cluster level. It is particularly important for capturing localized variations that would otherwise be obscured in aggregated analyses.

*Corresponding author: steve.adeshina@nilcuniversity.edu.ng

E. Validation Dataset

To evaluate classification performance, a validation dataset comprising 229 settlement clusters was constructed. This dataset includes 140 electrified clusters (Table 2), consisting of 130 communities obtained from Rural Electrification Agency (REA) mini-grid records and 10 field-verified grid-electrified communities. Additionally, 89 clusters were confirmed as unelectrified through field verification all with their coordinates. The inclusion of both electrified and unelectrified clusters ensures a balanced and representative evaluation framework, enabling robust assessment of classification performance across diverse real-world conditions. Ground-truth validation is particularly critical for distinguishing low-intensity electrification signals from background radiance noise.

Table 2 Validation Data set

Dataset	Description	Source
130 Clusters	REA functioning Minigrids	REA 2025
10 clusters	Visited and verified as electrified via grid extention	Field Survey
89 Clusters	Visited and verified as unelectrified	Field Survey

F. Performance Assessment & Optimal Threshold Selection

A confusion matrix was used to compare predicted electrification classifications against the ground-truth validation dataset at each candidate radiance threshold. Measures of performance in terms of precision and recall and the F1-score as per Section II.D are determined. F1-score was chosen as the main measure of evaluation since it can balance the precision and recall. This is especially significant in this research, as the dataset used has an imbalance of classes of

electrified and unelectrified settlements, and as such, it may be prejudiced in the more traditional accuracy measures. The best radiance value was selected as the radiance value which maximised the F1-score and yet had low false-positive rate. This makes sure that electrified settlements are properly acted without exaggerating electrification in low radiance regions. Through such a data-driven optimization strategy the classification scheme attains a more accurate and objective choice of threshold than a fixed threshold or heuristic choice. The outcome of this process is spatially explicit national electrification mark at the level of community clusters.

IV. RESULTS AND DISCUSSION

A. Radiance Threshold Evaluation

The distribution of VIIRS DNB 2024 annual radiance values across the settlement clusters exhibits a strongly right-skewed and heavy-tailed pattern, indicating that the majority of clusters are characterised by low radiance intensity while a relatively small proportion contributes to high radiance values. Approximately 84.92% of clusters record radiance values below $1 \text{ nW}\cdot\text{cm}^{-2}\cdot\text{sr}^{-1}$, highlighting the predominance of low-light or non-electrified conditions across rural Nigeria. This distribution reflects the underlying electrification structure, where most settlements are either unelectrified or weakly electrified, while a limited number of clusters associated with grid-connected regions contribute disproportionately to the high-radiance tail. To evaluate electrification classification performance, six candidate radiance thresholds ($0.25\text{--}0.50 \text{ nW}\cdot\text{cm}^{-2}\cdot\text{sr}^{-1}$) were assessed using a validation dataset comprising 229 settlement clusters (140 electrified and 89 unelectrified). Classification performance was quantified using confusion matrix metrics, including precision, recall, specificity, accuracy, and F1-score, as presented in Table 3.

Table 3 Confusion Matrix and Classification Performance Across Candidate Radiance Thresholds

Threshold ($\text{nW}\cdot\text{cm}^{-2}\cdot\text{sr}^{-1}$)	TP	FN	TN	FP	Precision	Recall	Specificity	Accuracy	F1-score
0.25	13	127	86	3	0.81	0.09	0.97	0.43	0.17
0.30	126	14	3	86	0.59	0.90	0.03	0.56	0.72
0.35	121	19	35	54	0.69	0.86	0.39	0.68	0.77
0.40	121	19	35	54	0.69	0.86	0.39	0.68	0.77
0.45	120	20	30	59	0.67	0.86	0.34	0.66	0.75
0.50	126	14	2	87	0.59	0.90	0.02	0.56	0.71

The results reveal a clear trade-off between precision and recall across the evaluated thresholds. At the lowest threshold ($0.25 \text{ nW}\cdot\text{cm}^{-2}\cdot\text{sr}^{-1}$), the model achieves high precision (0.81), indicating that predicted electrified clusters are largely correct; however, recall is extremely low (0.09), meaning that the majority of electrified settlements are missed. This demonstrates that overly conservative thresholds are unsuitable for detecting electrification in low-radiance rural environments.

As the threshold increases, recall improves substantially, reaching 0.90 at thresholds of 0.30 and $0.50 \text{ nW}\cdot\text{cm}^{-2}\cdot\text{sr}^{-1}$. However, this improvement is accompanied by a sharp decline in specificity, indicating a significant

increase in false-positive classifications. This implies that higher thresholds tend to overestimate electrification by incorrectly classifying unelectrified settlements as electrified. Among the tested thresholds, $0.35 \text{ nW}\cdot\text{cm}^{-2}\cdot\text{sr}^{-1}$ yields the highest F1-score (0.77), representing the best balance between precision (0.69) and recall (0.86). Although a similar performance is observed at $0.40 \text{ nW}\cdot\text{cm}^{-2}\cdot\text{sr}^{-1}$, the lower threshold (0.35) was selected as the optimal value to enhance sensitivity to low-intensity electrification signals typical of rural settlements. This result confirms the importance of threshold optimization, as fixed or arbitrarily selected thresholds can lead to either underestimation or overestimation of electrification coverage.

Figure 3 further illustrates the variation in classification performance metrics across the evaluated thresholds,

reinforcing the trade-off between detection sensitivity and classification reliability.

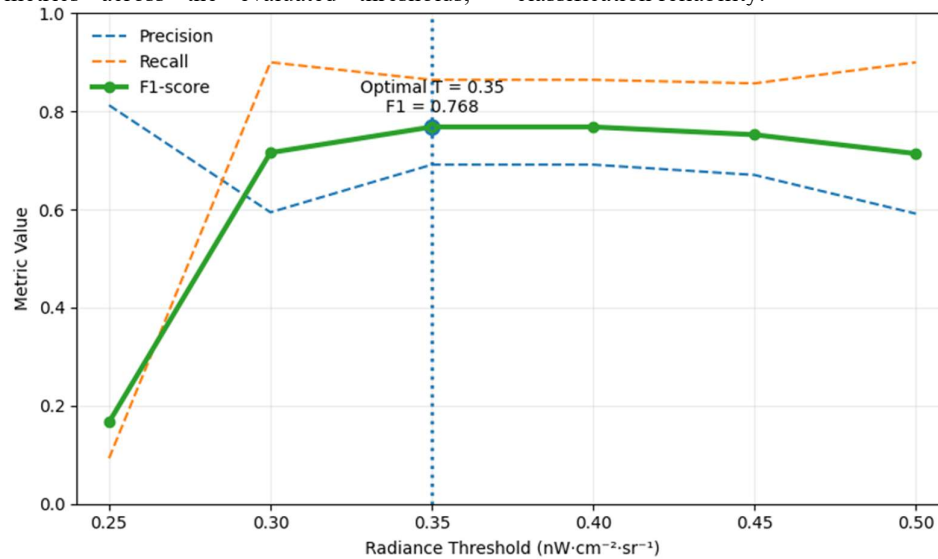


Figure 3 Classification performance of the VIIRS-based electrification classifier across evaluated radiance thresholds (0.25 – 0.50 $nW \cdot cm^{-2} \cdot sr^{-1}$).

B. National Electrification Updated Status

Applying the optimized threshold ($T^* = 0.35$ $nW \cdot cm^{-2} \cdot sr^{-1}$) to the full national dataset resulted in a binary classification of all 22,696 settlement clusters. The results indicate that 18,029 clusters (79.45%) are classified as unelectrified, while 4,667 clusters (20.55%) are classified as electrified. These findings reveal a substantial electrification gap at the community level, with approximately four out of every five rural settlement clusters lacking detectable

electrification. The classification effectively partitions the national radiance distribution into two distinct regimes: a dominant low-radiance regime corresponding to unelectrified or weakly electrified settlements, and a smaller high-radiance regime associated with electrified communities. The resulting high-resolution electrification map Figure 4 provides a spatially explicit representation of access disparities and serves as a foundational layer for identifying priority areas for electrification intervention.

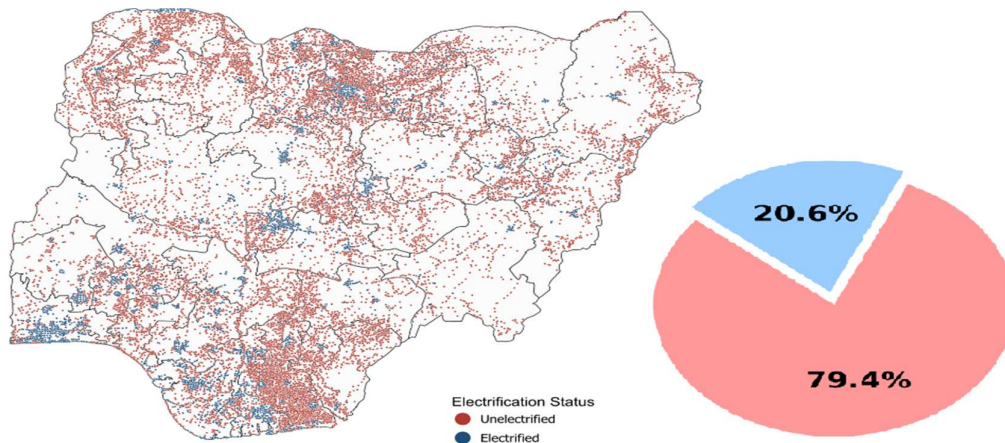


Figure 4 National Electrification Classification Map

C. Spatial Organization and Area Distribution.

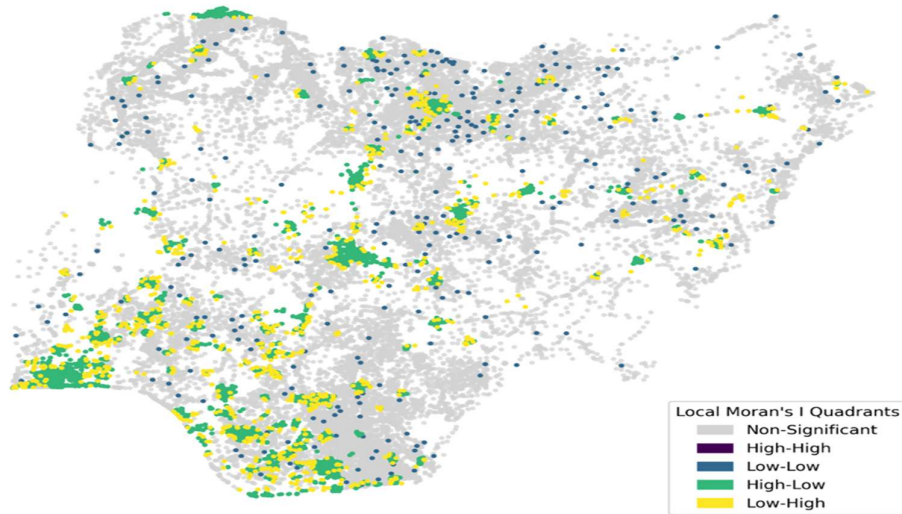
Spatial autocorrelation analysis was conducted to assess the geographic structure of electrification patterns. The Global Moran's I value of 0.5318 ($p < 0.001$) indicates strong positive spatial autocorrelation, confirming that electrification is not randomly distributed but instead exhibits significant clustering. This pattern is consistent across different spatial scales, as shown in Table 4, where Moran's I values remain consistently high for $k = 5, 8,$ and 10 nearest neighbors. The

high z-scores and statistically significant p-values further confirm the robustness of the observed clustering. Local Indicators of Spatial Association (LISA) analysis (Figure 5) reveals distinct spatial patterns, with electrified clusters concentrated along existing infrastructure corridors and unelectrified clusters predominantly located in remote rural regions. The presence of spatial outliers electrified clusters within predominantly unelectrified regions and vice versa, suggests ongoing but uneven electrification expansion.

Table 4 Global Moran's I Results (k = 5, 8, 10)

k (Nearest Neighbours)	Moran's I	Expected I	z-score	p-value
5	0.5713	-0.00004	144.39	< 0.001
8	0.5318	-0.00004	168.91	< 0.001
10	0.5148	-0.00004	182.37	< 0.001

Zonal analysis (Table 5) shows that the North-West and North-East regions contain the highest number of unelectrified clusters, both in absolute and relative terms, reflecting significant infrastructure deficits. In contrast, the South-West exhibits the lowest proportion of unelectrified clusters, indicating relatively higher electrification coverage.

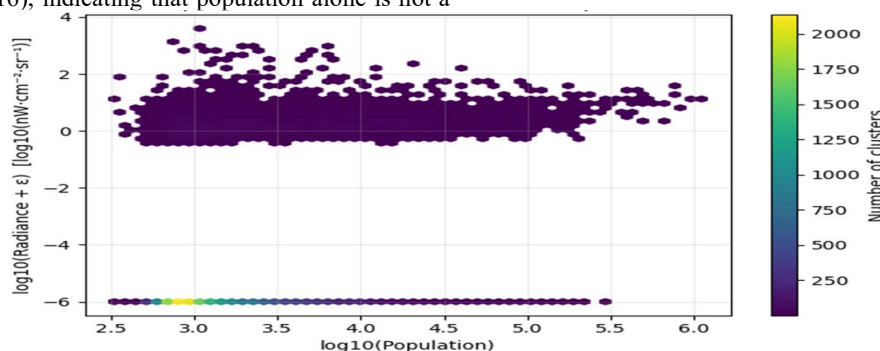
**Figure 5** Local Indicators of Spatial Association (LISA) Cluster Map**Table 5** Counts and shares of least-cost mode by zone

Zone	Unelectrified	Electrified	National%
North-West	6211	1066	34.45
North-East	2984	518	16.55
North-Central	2759	902	15.30
South-South	2648	1076	14.69
South-East	2050	261	11.37
South-West	1377	844	7.64

D. Relationship Between Radiance and Population

Correlation analysis was conducted to examine the relationship between radiance intensity and settlement population (log-transformed). The results indicate a moderate positive correlation ($Pearson\ r = 0.398, p < 0.001$), suggesting that larger settlements tend to exhibit higher radiance values. However, the explanatory power remains limited ($R^2 \approx 0.16$), indicating that population alone is not a

strong predictor of electrification status. Significant variation in radiance values is observed among settlements with similar population sizes, reflecting heterogeneous infrastructure distribution. This finding highlights the limitation of using demographic indicators as proxies for electrification and reinforces the importance of spatially explicit, radiance-based classification methods. Figure 6

**Figure 6** Relationship between VIIRS Radiance and Population

E. Implications for Electrification Planning

The results demonstrate that electrification patterns in Nigeria are highly structured and strongly influenced by existing infrastructure networks. Electrified settlements tend to form corridor like patterns aligned with grid infrastructure, while unelectrified clusters are concentrated in rural interior regions. This spatial configuration reflects a historical pattern of infrastructure-driven electrification expansion, where access has propagated outward from urban centres along transmission corridors. Consequently, large concentrations of unelectrified settlements particularly in northern regions represent critical priority areas for intervention.

In such contexts, decentralized electrification solutions, including mini-grids and standalone systems, provide viable pathways for extending access to dispersed communities. The high-resolution electrification dataset developed in this study provides a robust empirical foundation for national planning, enabling targeted deployment of electrification strategies and supporting subsequent techno-economic optimization. Overall, the proposed classification framework establishes a critical baseline for identifying unelectrified communities and serves as a foundational input for future modelling of energy demand, system design, and technology selection in rural electrification planning.

V. CONCLUSION

This study developed a cluster-based electrification classification framework for Nigeria using VIIRS night-time light radiance data integrated with high-resolution settlement cluster datasets (Husein et al., 2024). By applying a statistically optimized thresholding approach, the study enabled electrification assessment at the community level, providing a more spatially explicit and planning relevant alternative to conventional administrative level analyses. The results reveal a substantial electrification deficit across rural Nigeria, with 18,029 out of 22,696 settlement clusters (79.45%) classified as unelectrified. The radiance distribution was strongly right-skewed, reflecting the dominance of low-luminosity settlements and the limited spatial extent of electrified areas. Threshold evaluation demonstrated that an optimal radiance threshold of $0.35 \text{ nW}\cdot\text{cm}^{-2}\cdot\text{sr}^{-1}$ provides the best balance between precision and recall, highlighting the importance of data-driven threshold selection in improving classification reliability.

Spatial analysis further showed that electrification patterns are highly structured, with strong positive spatial autocorrelation (Moran's $I = 0.5318$, $p < 0.001$). Electrified settlements were observed to follow corridor-like patterns aligned with existing grid infrastructure, while unelectrified clusters were concentrated in remote rural regions. This reflects a historically infrastructure-driven electrification process, resulting in significant regional disparities in electricity access. The findings underscore the need for targeted and decentralized electrification strategies, particularly in regions with high concentrations of unelectrified settlements. Mini-grids and standalone systems offer viable solutions for extending access to dispersed rural communities. The high-resolution electrification dataset developed in this study provides a robust empirical foundation for national electrification planning and supports more effective resource allocation and policy design. Beyond its immediate application, this

work establishes a critical baseline for subsequent techno-economic analysis, including system sizing, technology selection, and cost optimization for rural electrification. Future research will extend this framework by integrating energy demand modelling, temporal variability in night-time light data, and advanced optimization techniques to support scalable and sustainable electrification planning.

AUTHOR CONTRIBUTIONS

Y. A. Badamasi: conceptualization, methodology, data curation, validation, formal analysis, and writing - original draft. **S. Adeshina:** supervision, conceptualization, and writing - review and editing. **A. D. Usman:** methodology review - review and editing. All authors read and approved the final manuscript.

REFERENCES

- Adelabu, A. S. and Olusola, A. (2021).** Mapping built-up areas using Sentinel-2 imagery. *IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, pp. 4484–4487.
- Coesfeld, J., Kuester, T., Kuechly, H. and Kyba, C. (2020).** Reducing variability in VIIRS night-time light data. *Sensors*, 20, pp. 1–20.
- Doll, C. and Pachauri, S. (2010).** Estimating rural populations without access to electricity in developing countries through night-time light satellite imagery. *Energy Policy*, 38, pp. 5661–5670.
- Earth Observation Group (2024a).** VIIRS Nighttime Light Data. Available at: <https://eogdata.mines.edu/products/vn/> (Accessed: 17 March 2026).
- Earth Observation Group (2024b).** VIIRS Nightfire Data. Available at: <https://eogdata.mines.edu/products/vnf/> (Accessed: 17 March 2026).
- Elvidge, C. D., Baugh, K., Zhizhin, M., Hsu, F. and Ghosh, T. (2017).** VIIRS night-time lights. *International Journal of Remote Sensing*, 38, pp. 5860–5879.
- Elvidge, C. D., Zhizhin, M., Ghosh, T., Hsu, F. and Taneja, J. (2021).** Annual time series of global VIIRS night-time lights derived from monthly averages. *Remote Sensing*, 13, pp. 1–14.
- Falchetta, G., Pachauri, S., Byers, E., Danylo, O. and Parkinson, S. (2019).** Satellite-based electrification dataset for Sub-Saharan Africa. *Scientific Data*, 6, pp. 1–9.
- Falchetta, G., Pachauri, S., Byers, E., Danylo, O. and Parkinson, S. C. (2020).** Global distribution of access to electricity. *One Earth*, 2, pp. 364–379.
- Gao, X., Wu, M., Niu, Z. and Chen, F. (2022).** Global identification of electrified areas using VIIRS data. *Remote Sensing*, 14, pp. 1–15.
- Husein, M., Moner-Girona, M., Falchetta, G., Stevanato, N., Fahl, F. and Szabó, S. (2024).** Settlement clustering for electrification planning. *Heliyon*, 10, e27440.
- IEA (2023).** World Energy Outlook 2023. Available at: <https://www.iea.org/reports/world-energy-outlook-2023> (Accessed: 17 March 2026).
- IEA, IRENA, UNSD, World Bank and WHO (2024).** Tracking SDG7: The Energy Progress Report 2024. Washington, DC: World Bank.

*Corresponding author: steve.adeshina@nileuniversity.edu.ng

Khavari, B., Korkovelos, A., Sahlberg, A., Howells, M. and Fuso Nerini, F. (2021). Spatial electrification modelling. *Scientific Data*, 8, pp. 1–17.

Li, X., Zhou, Y., Zhao, M. and Zhao, X. (2020). Mapping global urban extent using night-time lights. *Scientific Data*, 7, pp. 1–12.

Min, B., Gaba, K. M., Sarr, O. F. and Agalassou, A. (2013). Detection of electrification using night-time lights. *International Journal of Remote Sensing*, 34, pp. 8118–8141.

Wang, Z., Roman, M., Kalb, V., Miller, S., Zhang, J. and Shrestha, R. (2021). Assessment of VIIRS night-time light variability. *Remote Sensing of Environment*, 263, 112557.

World Bank (2024). Methodology for determining gas flare volumes from satellite data. Washington, DC: World Bank.

Yuan, X., Jia, L., Menenti, M. and Chen, Q. (2019). Improved detection of built-up areas using night-time lights. *Remote Sensing*, 11, 3002.

Yuan, X., Jia, L., Menenti, M. and Jiang, M. (2022). Night-time light data for urban analysis. *Big Earth Data*, 6, pp. 603–632.

Zhizhin, M., Elvidge, C. D., Ghosh, T., Gleason, G. and Bazilian, M. (2026). Global gas flaring monitoring using VIIRS Nightfire. *Remote Sensing*, 18, 314.