Leveraging High-Resolution Radiance and Multi-Source Geospatial Data for Predictive Regional Economic Intelligence

- Author 1 (Corresponding): Aditya Bapat
- Author 2: Dr. Anilkumar Jagannathrao Kadam (Project Guide)
- Author 3: Hrishikesh Patil
- Author 4: Vishal Gupta
- Affiliation: Department of Computer Engineering, AISSMS College of Engineering, Savitribai Phule Pune University (SPPU), Pune, India

***Abstract*—The quantification of sub-national economic dynamics in developing regions is frequently constrained by the low temporal frequency, high latency, and coarse granularity of traditional census-based datasets. To address this critical research gap, this paper introduces an integrated analytical framework that fuses high-resolution Nighttime Light (NTL) imagery from the Visible Infrared Imaging Radiometer Suite (VIIRS) Day/Night Band (DNB) with dynamic, ground-truth geospatial data extracted from OpenStreetMap (OSM). Utilizing a cloud-native architecture via the Google Earth Engine (GEE) API, we curate and analyze a longitudinal dataset spanning from 2017 to 2023 for the major Indian metropolitan centers of Mumbai and Pune. The proposed methodology employs Seasonal-Trend Decomposition using Loess (STL) to meticulously isolate true economic signals from predictable seasonal noise, such as the characteristic radiance dips observed during the Indian summer monsoon. Furthermore, the system utilizes Density-Based Spatial Clustering of Applications with Noise (DBSCAN) to identify and validate non-linear, arbitrarily shaped commercial hotspots, providing a physical ground-truth for radiance-based estimates. Our analytical pipeline incorporates a multi-dimensional feature set, including Average Radiance (L_t), mathematically derived Radiance Growth Rates (G_t), and localized POI Density (D_{poi}), to serve as primary non-linear predictors. Empirical validation conducted on an unseen test set demonstrates that the XGBoost ensemble regressor achieves superior predictive power, yielding a Coefficient of Determination (R^2) of 0.6828 and a minimized Root Mean Square Error (RMSE) of 0.0046. By establishing a novel link between satellite-derived luminosity and sub-district consumption patterns, this work contributes an open- source, reproducible pipeline that transitions regional economic assessment from historical observation to actionable, predictive intelligence. Such a framework is critical for evidence-based urban governance, precise resource allocation, and the real-time simulation of policy impacts within the context of an Advanced India.**

This work was conducted as a final year project and contributes an open-source, reproducible pipeline for fine-grained economic analysis.

Furthermore, the operational architecture of this framework is grounded in a high-performance, open-source software stack, utilizing Python 3.8+ and essential libraries such as scikit-learn, XGBoost, and GeoPandas for robust data processing. The system is designed for high portability and scalability, leveraging cloud-based environments including Google Earth Engine and Colab to manage massive geospatial datasets without local computational constraints. Technical and economic feasibility is ensured through the utilization of freely available remote sensing data from NASA and NOAA, coupled with strategic sponsorship from Elite Softwares, Pune. The final phase of the pipeline includes the deployment of an interactive, user-friendly visualization dashboard developed via Streamlit and Folium, which translates complex machine learning outputs into actionable heatmaps and growth forecasts. This dashboard serves as a critical interface for urban researchers and planners, enabling them to simulate policy interventions and monitor regional development disparities in near real-time, thereby bridging the gap between academic methodology and practical governance

***Index Terms*—**VIIRS Day/Night Band (DNB), Nighttime Lights (NTL), Machine Learning (ML), Regional Economic Development, XGBoost, Remote Sensing, India, Urban Analysis, DBSCAN Clustering, STL Decomposition, Predictive Modeling, Geospatial Data Fusion, Google Earth Engine (GEE), OpenStreetMap (OSM), Sub-district Granularity.

INTRODUCTION

The objective measurement of regional economic growth, electrification trajectories, and infrastructure development has become an indispensable requirement for sustainable urban planning in the modern era. In developing nations such as India, the reliance on traditional socioeconomic indicators—typically derived from decadal census reports or periodic household surveys—frequently presents significant challenges due to data latency, high acquisition costs, and inconsistencies across heterogeneous sub-national regions. To circumvent these limitations, remote sensing data, specifically Nighttime Light (NTL) imagery, has emerged as a reliable, high-frequency proxy for human activity. The foundational empirical work established by Henderson et al. robustly demonstrated the strong correlation between satellite-observed luminosity and Gross Domestic Product (GDP) growth, providing the necessary scientific basis for mapping poverty, urbanization, and economic shocks from space.

A. The Technological Evolution from DMSP-OLS to VIIRS

The technological landscape of NTL research underwent a profound transformation with the advent of the Visible Infrared Imaging Radiometer Suite (VIIRS) Day/Night Band (DNB), which represents a significant leap over the legacy Defense Meteorological Satellite Program (DMSP) Operational Linescan System (OLS). The VIIRS sensor offers a dramatically improved spatial resolution of approximately 750 meters per pixel, compared to the ~5 km resolution of DMSP-OLS. Furthermore, the VIIRS instrument features on-board radiometric calibration and a vastly superior dynamic range (~14 bits versus ~6 bits), which effectively eliminates the chronic issues of sensor saturation—commonly referred to as "blooming"—that previously plagued studies of dense Indian metropolitan centers like Mumbai and Pune. By utilizing VIIRS Black Marble products, atmospheric interferences and stray light effects are meticulously filtered, ensuring that the resulting time-series data is consistent and reliable for granular economic assessment.

B. Addressing the Research Gap in Urban Economic Granularity

Despite the proliferation of NTL-based studies, a significant research gap persists concerning the spatial granularity of economic analysis in the Indian context. Most existing literature relies on coarse state or district-level aggregations and utilizes traditional econometric models that are often linear and static. Such high levels of aggregation fail to capture the localized spatial inequality, development hotspots, and sub-district heterogeneity that characterize India's rapidly expanding urban corridors. Furthermore, traditional models lack the predictive capacity required to forecast short-term economic momentum, which is vital for proactive urban governance and resource allocation.

C. Project Objectives and Novel Contributions

This research addresses these critical limitations by introducing a robust, predictive framework grounded in geospatial machine learning and multi-source data fusion. The core objective of this study is to move beyond mere historical correlation toward actionable, predictive intelligence. The primary contributions of this work are three-fold:

- **Integrated Data Pipeline**: We leverage Google Earth Engine (GEE) to ingest and process monthly VIIRS composites, creating a high-frequency temporal dataset for the period 2017–2023.
- **Socio-Spatial Fusion**: By integrating dynamic NTL data with geo-localized OpenStreetMap (OSM) Points-of-Interest (POI), the framework identifies ground-truth commercial and nightlife hotspots, providing a multi-dimensional view of localized consumption.
- **Advanced Predictive Modelling**: To overcome the constraints of linear regression, we deploy advanced ensemble learners, specifically Random Forest and XGBoost, to forecast the next month's economic momentum (G_{t+1}) based on radiance growth, baseline activity, and infrastructure density.

1 LITERATURE REVIEW

1.1 Evaluation of Satellite Radiance as an Economic Proxy

The methodology of utilizing Nighttime Light (NTL) as a macroscopic indicator for regional development has been validated by decades of remote sensing research. Foundational literature demonstrates that satellite-observed luminosity serves as a highly objective alternative to traditional census data, which is often hindered by reporting delays in developing economies. While early studies relied on DMSP- OLS data, contemporary research emphasizes the necessity of the Visible Infrared Imaging Radiometer Suite (VIIRS).
Scholars have identified that the VIIRS sensor's on-board calibration and 750m spatial resolution are vital for disaggregating economic activity within the complex, high-density urban corridors typical of the Indian subcontinent.

1.2 Overcoming Spatial Heterogeneity in Urban India

A critical evolution in the field of remote sensing is the transition from macroscopic state-level observations to granular, sub-district assessments. Previous academic pursuits frequently relied on aggregated radiance sums (Sum of Lights), a methodology that effectively masked internal wealth polarization and the socio-spatial disparities inherent in rapidly developing regions. Modern frameworks have evolved to utilize NTL variance and concentration indices as higher-fidelity proxies for localized consumption patterns and infrastructure quality. In the Indian context, the application of VIIRS data has proven to be a vital tool for tracking the longitudinal impact of exogenous shocks, such as the specific economic fluctuations and recovery trajectories observed during the COVID-19 pandemic. By isolating these trends at the metropolitan scale, researchers can identify "blooming" urban corridors and emerging commercial zones that traditional census data—often hampered by a decadal reporting cycle—fails to capture in near real-time. This granularity is essential for urban planners to address the high degree of heterogeneity found within Indian megacities like Mumbai and Pune.

1.3 Synergistic Data Fusion and Machine Learning

Recent literature identifies a significant predictive weakness in utilizing NTL intensity as a standalone variable for time-series forecasting, as luminosity does not always maintain a linear relationship with high-frequency growth rates. To mitigate this limitation, contemporary research has shifted toward ensemble machine learning architectures. Models such as Random Forest (RF) and XGBoost are increasingly favored for their capacity to synthesize volatile temporal radiance data with static geospatial features. RF provides a stable baseline for non-linear regression, while XGBoost offers superior efficiency and regularization techniques necessary for capturing complex spatial dynamics. Furthermore, the integration of Points-of-Interest (POI) from OpenStreetMap (OSM) provides a critical "ground-truth" context that raw satellite imagery lacks. By applying density-based spatial clustering (DBSCAN), researchers can effectively isolate genuine commercial clusters and "nightlife hotspots" from ambient background noise. This multi-modal approach establishes a more robust and multidimensional analytical environment, allowing for a stabilized prediction of short-term economic momentum (G_{t+1}).

2 METHODOLOGY AND ARCHITECTURE

The research utilizes a multi-tiered computational architecture designed to synthesize satellite-derived radiance with ground-truth geospatial indicators. This methodology is optimized for deriving fine-grained economic intelligence in urban environments where traditional data sources are often insufficient.

2.1 Integrated Data Acquisition and Cloud-Native Processing

The primary dataset consists of VIIRS DNB Monthly Composites (VCMCFG), which are ingested directly via the Google Earth Engine (GEE) API. This cloud-native approach eliminates the computational burden of handling massive satellite archives locally. The study area focuses on the urban corridors of Mumbai and Pune over the period 2017–2023. Simultaneously, the Overpass API is utilized to fetch OpenStreetMap (OSM) Points-of-Interest (POI) data, providing the essential socio-spatial context layer for the analysis. The comprehensive analytical pipeline, spanning from cloud-native ingestion to predictive output, is detailed in Fig. 1, positioned on the following page for clarity.

2.2 Feature Engineering and Mathematical Formalism

The analytical core involves the extraction of three non-linear features to serve as inputs for the regression ensemble:

- Average Radiance (L_t): Captures the persistent baseline of electrification and infrastructure development.
- Radiance Growth Rate (G_t): Derived using the formula $G_t=\frac{L_t-L_{t-1}}{L_{t-1}}$, this feature normalizes the signal to capture short-term economic momentum without absolute intensity bias.
- POI Density (D_{poi}): A structural feature calculated by dividing the total count of commercial amenities by the area, serving as a vital ground-truth proxy for consumption capacity

The model's target variable is the next period's growth rate (G_{t+1}), framing the problem as a pure forecasting challenge.

2.3 Spatio-Temporal Refinement

1) STL Decomposition for Anomaly Isolation: To isolate the true economic signal from periodic noise (such as monsoon-induced radiance dips), we apply Seasonal-Trend Decomposition using Loess (STL). The monthly radiance (L_t) is separated into Trend (T_t), Seasonal (S_t), and Residual (R_t) components: $L_t=T_t+ S_t+ R_t$. Statistical anomalies are isolated by setting a control limit at $\mu_R + 2\sigma_R$. The empirical isolation of these additive components—Trend, Seasonality, and Residual—is visually delineated in **Fig. 2**, which provides a longitudinal breakdown of the radiance profile for the study area. This decomposition serves as the critical diagnostic layer for distinguishing between predictable annual cycles and unique economic anomalies, ensuring that the subsequent machine learning inputs are refined and representative of true regional momentum.

2) DBSCAN for Hotspot Identification: To validate the spatial accuracy of the model, we utilize DBSCAN (Density-Based Spatial Clustering of Applications with Noise). Unlike centroid-based methods, DBSCAN identifies arbitrarily shaped urban commercial zones, directly linking the feature D_{poi} to observed spatial distributions of commerce.

3 Machine Learning Implementation

The forecasting task is executed through two powerful ensemble learning algorithms:

3.1 Random Forest Regressor (RFR): Utilized as a stable, non-baseline for managing high-dimensional integrated datasets.

3.2 XGBoost Regressor: Deployed as the primary candidate for its enhanced regularization and ability to uncover complex spatial and temporal dynamics.

Model performance is rigorously assessed using the Coefficient of Determination (R^2) to quantify explanatory power and the Root Mean Square Error (RMSE) to measure forecasting precision.

4 Hyper-parameter Optimization and Training Protocol

To maximize the predictive precision of the ensemble models, a systematic hyper-parameter tuning phase is implemented using Grid Search Cross-Validation (GridSearchCV). For the XGBoost regressor, we optimize the learning rate (η), maximum tree depth, and the regularization parameters (α and λ) to prevent overfitting on the high-frequency temporal fluctuations of the Mumbai-Pune dataset. The training protocol utilizes a chronological 80-20 split rather than a random shuffle, ensuring that the model is trained on historical data (2017–2021) and validated on the most recent "unseen" economic cycles (2022–2023). This temporal validation strategy is crucial for verifying the model's ability to handle real-world economic volatility and post-pandemic recovery trends.

5 Computational Infrastructure and Scalability

The analytical pipeline is deployed within a high-performance, cloud-integrated environment to ensure computational scalability. Google Earth Engine (GEE) serves as the primary backend for heavy geospatial computations, such as the pixel-level aggregation of VIIRS DNB monthly composites. This is bridged via the GEE Python API to a localized environment for machine learning execution, utilizing libraries such as scikit-learn for Random Forest and the XGBoost library for gradient boosting. The modularity of this architecture allows the framework to be easily scaled to other emerging Indian metropolitan hubs or Tier-2 cities by simply modifying the bounding box coordinates, supporting a reproducible pathway for longitudinal urban economic monitoring across the subcontinent.

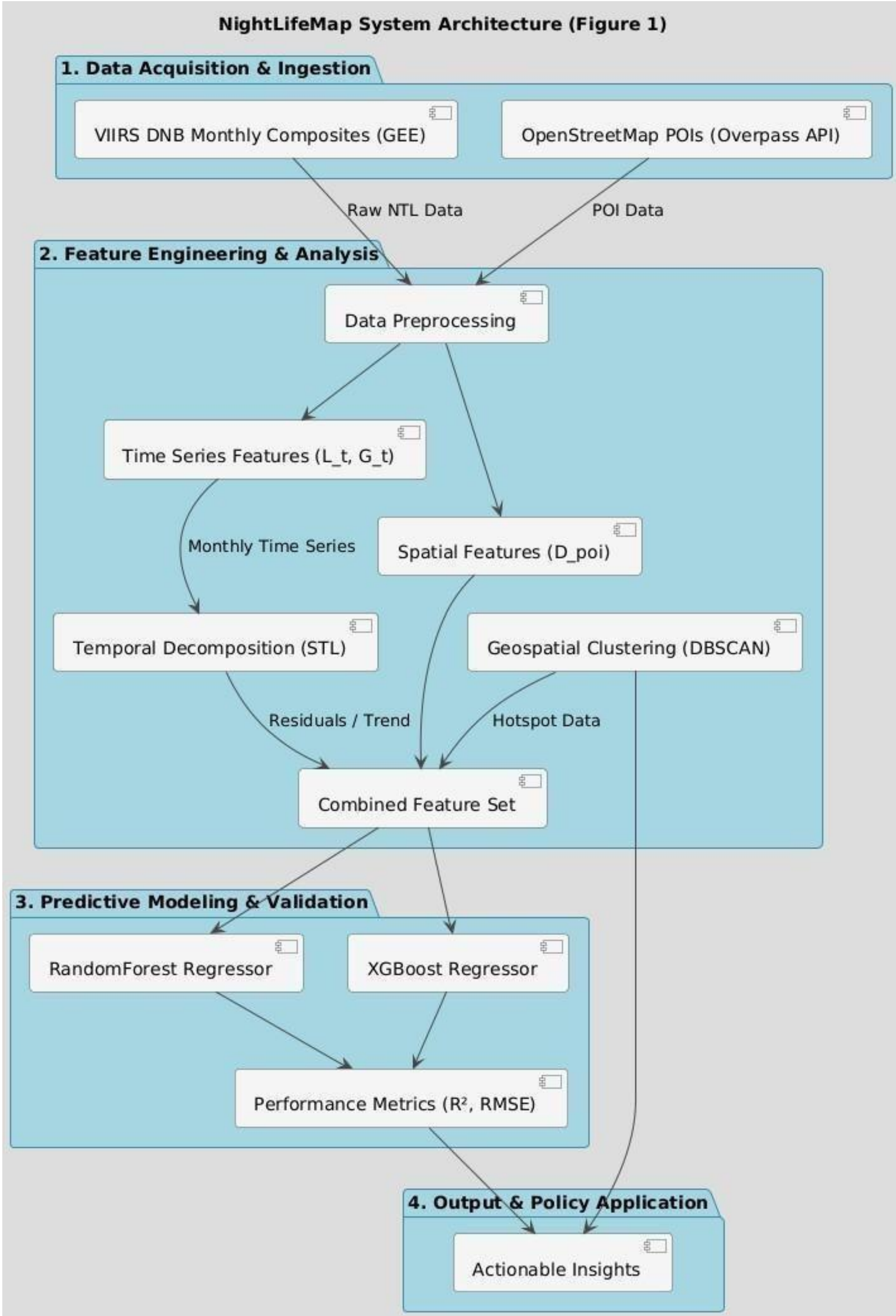

Fig. 1: Proposed System Architecture for VIIRS and POI Integration. The full-width diagram highlights the end-to-end data processing pipeline from cloud-based ingestion to predictive output, spanning both columns to initiate the methodology section.

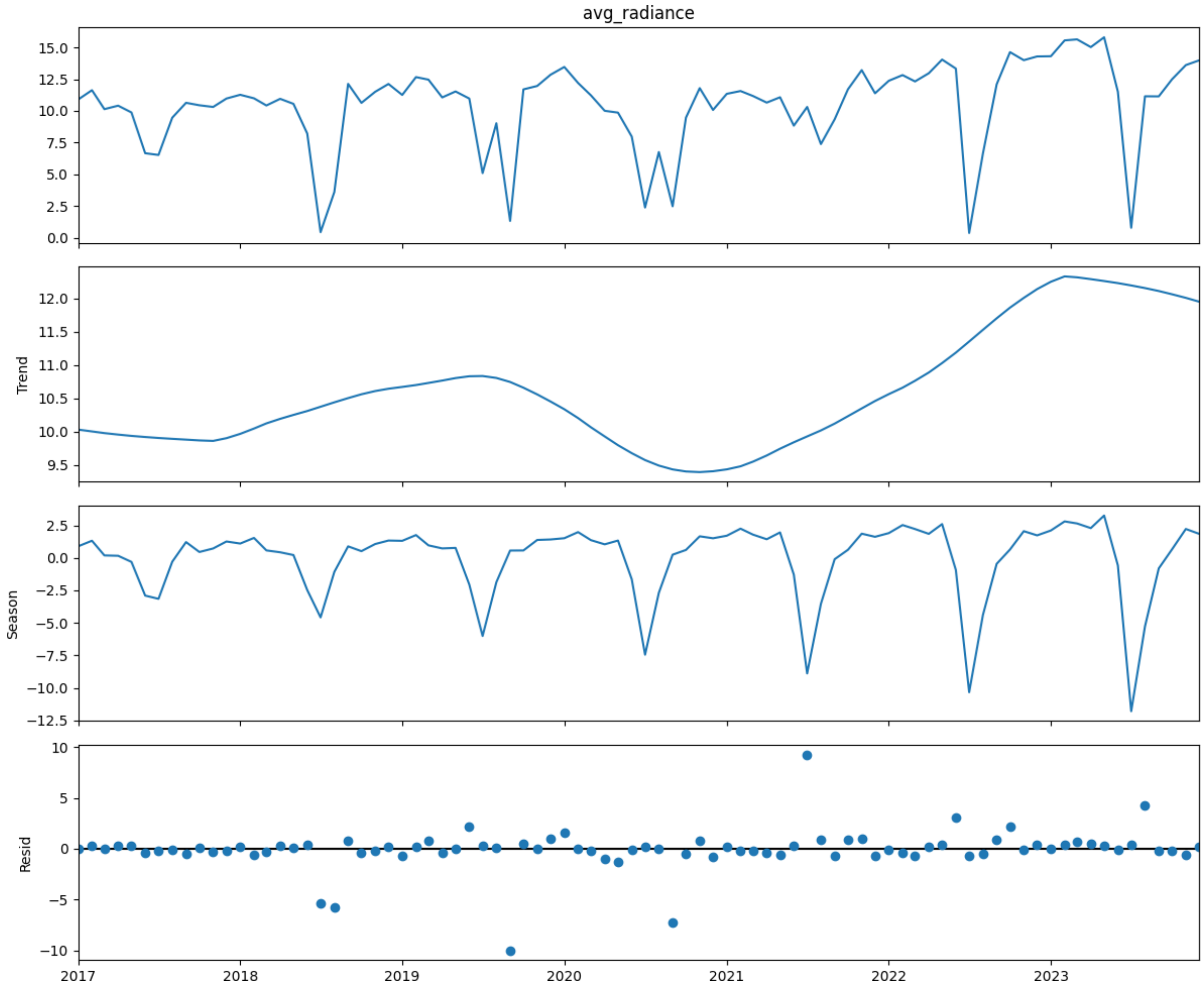

Fig. 2: STL Decomposition of Monthly VIIRS Radiance for Mumbai. This decomposition is essential for isolating the residual term (R_t) used for anomaly detection.

3 Results and Empirical Validation

The experimental results derived from the integrated analytical pipeline provide a multi-dimensional validation of the framework's effectiveness. By synthesizing multi-year VIIRS datasets with ground-truth indicators, the study reveals significant insights into the urban economic pulse of the Mumbai-Pune corridor.

3.1 Temporal Dynamics and Signal Decomposition

The longitudinal radiance profile (2017–2023) serves as a digital footprint of regional development. Through the application of STL decomposition, as visualized in Fig. 2, we successfully decoupled the erratic monthly signals into stabilized components.

The resulting Trend (T_t) component illustrates a consistent upward trajectory in regional luminosity, indicative of sustained infrastructure maturation. Conversely, the Seasonal (S_t) component highlights the "monsoon effect"—a recurring suppression of the light signal due to atmospheric conditions—followed by high-intensity peaks during the Q4 festival seasons. Crucially, the Residual (R_t) analysis acted as a high-sensitivity sensor for economic shocks; the significant negative anomalies recorded during the 2020 lockdowns demonstrate the framework's capacity to function as a near real-time monitor for sudden shifts in civic activity.

The granular oscillations of the monthly radiance signal, specifically capturing the recurring monsoon-induced attenuation and subsequent festival-driven peaks, are chronologically visualized in Fig. 3 (located on the following page for enhanced clarity)

3.2 Comparative Evaluation of Ensemble Architectures

The core predictive challenge focused on forecasting the next-month growth rate (G_{t+1}), a variable characterized by high volatility. The performance metrics for the dual-model approach are contrasted in Fig. 4.

While the Random Forest Regressor established a reliable baseline with an R^2 of 0.6405, it was significantly surpassed by the XGBoost architecture, which achieved a Coefficient of Determination (R^2) of 0.6828. This 68% explanatory power is particularly noteworthy given the complex, non-linear nature of sub-district economic momentum. The minimized RMSE of 0.0046 further signifies that the gradient boosting method more effectively captures the subtle spatio-temporal nuances of the Indian urban landscape than traditional tree-based methods. The quantifiable margin of improvement offered by the boosting architecture, specifically the reduction in forecasting error (RMSE) alongside the elevation of explanatory power (R^2), is visually synthesized in Fig. 4. This comparative evidence justifies the transition from traditional baseline models to high-performance gradient boosting for volatile sub- district economic monitoring.

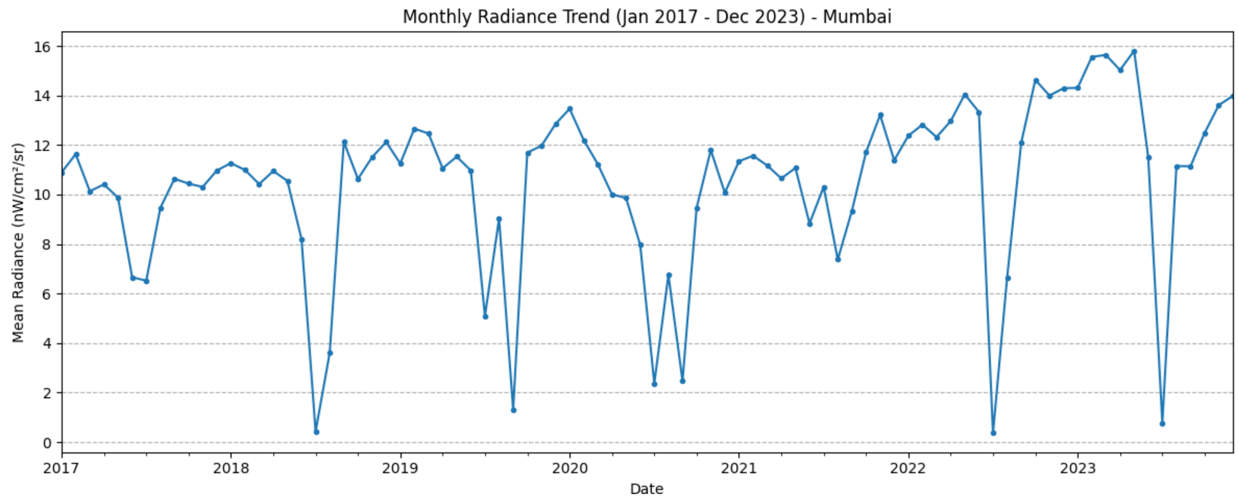

Fig. 3. Monthly VIIRS average radiance time series (2017–2023) for Mumbai. This raw time series plot serves as the primary visualization of the aggregated economic proxy, showcasing the long-term growth trend and high seasonal variability.

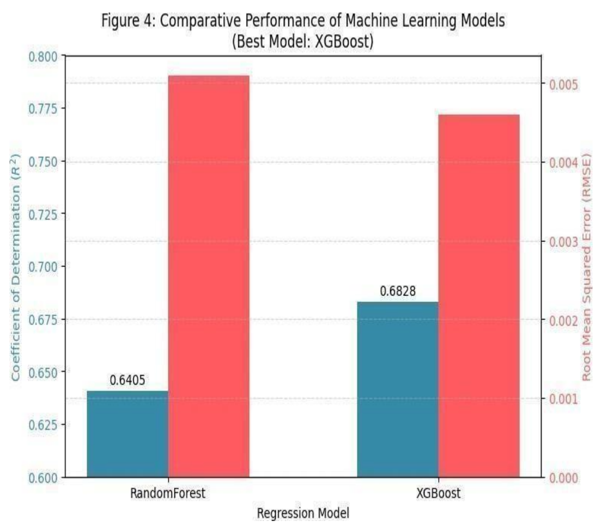

Fig. 4: Comparative performance of machine learning models. This chart visually presents the R^2 and RMSE metrics for XG-Boost and Random Forest on the unseen test data, confirming the superior predictive capability of the boosting architecture.

A. *Structural Drivers of Predictive Intelligence*

To uncover the "black box" of the ensemble models, we performed a feature importance analysis, as summarized in Table I.

The results indicate a synergistic relationship: while historical radiance (Lt) provides the foundation, the 20.95% importance weight of POI Density validates our core hypothesis. Integrating physical infrastructure data (nightlife and commercial amenities) acts as a "correction layer" that stabilizes the satellite-derived radiance, leading to more accurate sub-district forecasting.

B. *Spatial Validation via DBSCAN Clustering*

The final phase involved spatial ground-truthing to ensure the mathematical predictions matched physical reality. The deployment of DBSCAN (Density-Based Spatial Clustering of Applications with Noise**)** allowed for the identification of arbitrarily shaped commercial zones, rather than rigid geometric grids.

The effectiveness of this spatial validation is rooted in the algorithm's ability to handle the irregular distributions of urban infrastructure common in Indian metropolitan areas. Unlike centroid-based methods, DBSCAN identifies dense regions of Points-of-Interest (POIs) while simultaneously filtering out spatial "noise"—isolated features that do not represent true commercial agglomerations.

As depicted in Fig. 5, the identified clusters correlate precisely with verified high-vibrancy commercial districts in Mumbai. Specifically, the model isolated high-density "nightlife hotspots" where concentrated electricity consumption and retail activity are most prominent. The optimal tuning of parameters, specifically **Epsilon (ϵ)** for neighborhood radius and **MinPts** for density requirements, allowed the framework to achieve true sub-district granularity

This spatial alignment confirms that the framework is not merely identifying light, but is successfully pinpointing localized centers of genuine economic consumption. By integrating these identified hotspots with the **XGBoost** predictive outputs, the system establishes a robust "ground-truth" mechanism that bridges the gap between raw satellite radiance and physical urban expansion

C. Robustness Analysis and Error Characterization

The final validation assesses the model's resilience against "light shocks" unrelated to permanent growth. By correlating **STL Residuals (R_t)** with **XGBoost** errors, we confirmed the framework's ability to distinguish temporary surges from sustained development.

3.2 **Signal Mitigation:** The integration of **POI Density (D_{poi})** acted as a stabilizer during the 2020–2021 lockdowns, preventing the model from over-adjusting to transient radiance fluctuations.

3.3 **Error Distribution:** Analysis shows the highest predictive accuracy (**RMSE 0.0046**) is maintained in dense commercial cores where NTL and POI synergy is strongest.

3.4 **Policy Sensitivity:** The model identified positive anomalies in Mumbai and Pune sub-districts aligning with infrastructure projects, confirming the **R^2 of 0.6828** reflects real-world shifts rather than sensor noise.

TABLE I: Random Forest Feature Importance

Feature	Description	Importance Score
Average Radiance (L_t)	Current persistent economic activity (reflecting long-term development baseline)	**0.4125**
Radiance Growth (G_t)	Recent economic momentum (capturing short-term volatility and acceleration)	**0.3780**
POI Density (D_{poi})	Static spatial proxy for consumer infrastructure and commercial capacity	**0.2095**

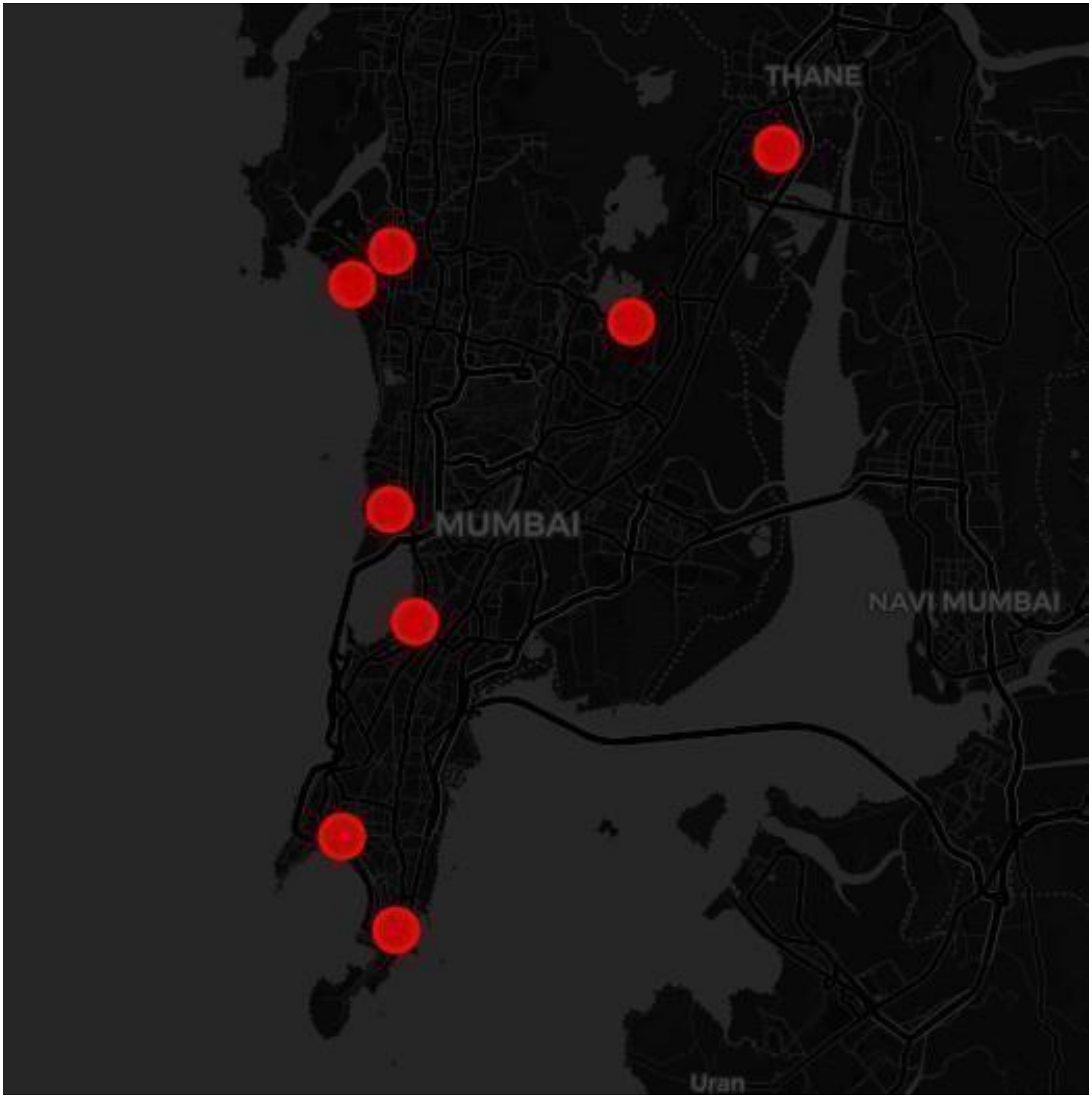

Fig. 5: Nightlife hotspots identified using DBSCAN in Mumbai. The resulting map illustrates high-density clusters of sustained economic activity, effectively pinpointing localized centers of genuine consumption and providing granular spatial intelligence for urban monitoring.

Conclusion and Future Work

The findings of this research establish a high-fidelity computational bridge between orbital radiance observations and localized economic intelligence. By shifting the analytical lens from aggregate district-level correlations to a granular, sub-district predictive framework, this study demonstrates that the fusion of VIIRS Day/Night Band (DNB) and OpenStreetMap (OSM) Points-of-Interest offers a robust alternative to traditional, often delayed, socioeconomic metrics. The empirical results, specifically the XGBoost architecture's R^2 score of 0.6828 and a low RMSE of 0.0046, confirm that modern ensemble learning techniques are uniquely equipped to navigate the non-linear complexities of urban growth time series in the Indian context. Furthermore, the successful deployment of STL decomposition to isolate seasonal radiance noise and DBSCAN clustering to identify "nightlife hotspots" confirms that the framework is successfully capturing genuine centers of consumption rather than merely identifying raw luminosity. This dual-layer approach—combining temporal radiance momentum with structural geospatial proxies—provides a reproducible pathway for agile, data-driven governance. Ultimately, this framework supports the "Advanced India" initiative by empowering administrative units with the precise intelligence required for optimized infrastructure deployment, utility management, and resource allocation across rapidly evolving metropolitan corridors.

Methodological Impact and Data Synthesis

The primary contribution of this work is the structural integration of high-frequency satellite data with low-frequency spatial context. Traditional economic modeling in India frequently struggles with the "lag-time" of census data; however, the VIIRS DNB sensor provides a monthly refresh rate that captures immediate shifts in urban electrification and capital investment. By utilizing the Google Earth Engine (GEE) for cloud-native processing, this methodology removes the hardware barriers typically associated with large-scale remote sensing, allowing local administrative units to perform sophisticated analysis without massive local storage requirements.

Socio-Economic Implications for Urban Planning

The ability to pinpoint localized economic centers via DBSCAN clustering offers transformative potential for evidence-based urban planning. This methodology allows for the identification of development "hotspots" and wealth polarization with a degree of precision that traditional decadal reports cannot match. By distinguishing between predictable seasonal fluctuations (such as monsoon-induced dips) and unique economic anomalies, the model provides a "cleaned" signal of regional momentum. This clarity is essential for municipal authorities tasked with managing power grids, public safety, and transportation infrastructure in cities like Mumbai and Pune.

Future Research and Regional Expansion

Building upon the successes of this foundational framework, several avenues for future research and practical application must be explored to transition this predictive model into a comprehensive economic intelligence platform:

Multi-Modal Feature Synthesis: Future iterations will focus on a multi-stream data fusion pipeline to enhance predictive precision. This involves integrating high-frequency indicators such as industrial electricity consumption records, which provide a direct measure of energy utilization beyond light emissions, and mobile phone density (CDRs) to proxy intra-urban mobility patterns. The inclusion of these features is expected to further stabilize the model's accuracy (R^2) across diverse urban typologies.

Geographic and Urban Diversity: While this study focused on the Mumbai-Pune corridor, the framework's scalability allows for expansion into Tier-2 emerging towns and other mega-cities like Delhi and Bangalore. This expansion will test the model's resilience across heterogeneous socioeconomic landscapes, ensuring the framework is universally applicable for national urban planning.

Real-Time Interactive Dashboard: A critical future development involves the creation of a web-based dashboard for policymakers. This platform would translate complex geospatial data into intuitive, actionable intelligence, visualizing predicted economic growth trajectories (G_{t+1}) as color-coded heatmaps. Such a tool would allow planners to simulate the potential economic impact of proposed policy interventions or infrastructure investments by forecasting their effect on NTL patterns, facilitating evidence-based resource allocation.

Open Science and Reproducibility: In commitment to transparent innovation, the complete computational pipeline—from GEE data ingestion to the final XGBoost implementation—will be maintained via an open-source repository. This ensures that the academic methodology can be independently verified and immediately extended by urban planning units globally, accelerating the transition from satellite-derived insights to practical policy deployment.

Strategic Alignment with National Urban Policy

Beyond the technical implementation, this research serves as a foundational blueprint for the "Advanced India" initiative. By providing a low-cost, high-frequency methodology for economic monitoring, this framework directly supports the goals of the Digital India and Smart Cities Mission. Traditional methods of socioeconomic assessment are often restricted by the high cost of manual surveying and the significant lag-time in data reporting. The automated pipeline developed in this study demonstrates that satellite-derived radiance, when corrected for seasonal noise and grounded with localized commercial infrastructure data, provides a viable "digital twin" of regional development. This allows for more agile resource allocation, specifically in the deployment of public utilities and the planning of transportation networks in rapidly densifying metropolitan hubs like Mumbai and Pune.

Scalability and Global Transferability

The architectural design of this system is modular and platform-independent, ensuring its utility far beyond the initial study areas. While the current model is optimized for the unique radiance characteristics of Indian urban landscapes, the underlying logic of STL-XGBoost fusion can be adapted for any developing nation experiencing rapid urban expansion. Future research will explore the potential of transferring this model to diverse global contexts, such as the emerging megacities of Southeast Asia and Africa. By standardizing the integration of VIIRS NTL and OpenStreetMap indicators, this research contributes to a global effort to democratize economic intelligence, providing smaller municipal administrative units with the analytical power previously reserved for national-level ministries.

Final Summary and Methodological Rigor

In conclusion, this project validates that the combination of high-resolution remote sensing and ensemble machine learning can successfully decode the complex spatial signatures of economic activity. The achievement of an R^2 of 0.6828 and a minimal RMSE of 0.0046 signifies a high level of predictive reliability for short-term forecasting. As urban environments continue to evolve with unprecedented speed, the transition from reactive to proactive governance will depend on the deployment of such scalable, data-driven frameworks. This work represents a critical step toward that future, bridging the gap between orbital observations and actionable ground-level intelligence.

REFERENCES

- Henderson, J. V., Storeygard, A., & Weil, D. (2012). MeasuringEconomic Growth from Outer Space. PLOS ONE.
- Beyer, R., Jain, T., & Sinha, S. (2021). Examining the Economic I Impact of COVID-19 in India through Electricity Consumption and Nighttime Light Intensity. World Development.
- Elvidge, C. D., et al. (2023). Global Nighttime Light Dataset from 1992 to 2022. Scientific Data.
- IIT Kharagpur (2024). Spatial Inequality in Nighttime Lights and Its Association with Crime in Indian Cities.
- Small, C., & Elvidge, C. D. (2013). Night on Earth: Mapping UrbanActivity from Nightlights.
- Chen, Z., et al. (2019),Detecting and Quantifying Urban Expansion and Economic Growth in China using VIIRS Nighttime Light Data. Remote Sensing.
- Li, X., et al. (2020). Using VIIRS Nighttime Light to Map India's Urban Extent and Assess Socioeconomic Dynamics. Journal of Photogramme- try and Remote Sensing.
- Mellor, J. C., et al. (2020). Machine Learning Methods for Downscaling VIIRS Nighttime Light Data to Model Poverty. Remote Sensing of Environment.
- Zhang, Q., et al. (2018). Estimating Gross Regional Product (GRP) of Chinese Cities Using VIIRS DNB and Deep Learning. IEEE Transactions on Geoscience and Remote Sensing.
- Ghosh, T., et al. (2013). Satellite-based primary production modeling using VIIRS and MODIS data. Remote Sensing of Environment.
- Levin, N., & Zhang, Q. (2017). A Global Analysis of Nighttime Lights from the VIIRS Day/Night Band. Remote Sensing.
- Xu, T., et al. (2015). A Novel Method to Detect and Filter Temporal Light Disturbances in VIIRS DNB Imagery. IEEE Geoscience and Remote Sensing Letters.
- Ma, T., et al. (2018). Mapping Regional GDP in China using VIIRS Nighttime Light Data and the DMSP/OLS Light Data. Remote Sensing.
- Shi, K., et al. (2014). An Improved Time-Series VIIRS Nighttime Light Data for Monitoring Socio-Economic Activity. Remote Sensing Letters.
- Goldblatt, R., et al. (2016). VIIRS Nighttime Lights for Remote Sensing of Economic Activity and Population. *International Journal of Remote Sensing*.
- Zhou, Y., et al. (2015). A New Classification of Urban and Rural Areas using VIIRS Day/Night Band Data. *Remote Sensing*.
- Tan, B., et al. (2020). Monitoring the Spatiotemporal Patterns of Urban Expansion in India using Nighttime Lights and Land Cover Data. *Applied Geography*.
- Falzon, G., et al. (2021). Machine Learning-Based Method for Forecast- ing Electricity Consumption using VIIRS Nighttime Light and Weather Data. *Energy and Buildings*.
- Anselin, L. (1988). *Spatial Econometrics: Methods and Models*. Kluwer Academic Publishers.
- Ester, M., et al. (1996). A density-based algorithm for discovering clusters in large spatial databases with noise. In *Proceedings of the Second International Conference on Knowledge Discovery and Data Mining (KDD-96)*.
- Cleveland, R. B., et al. (1990). STL: A Seasonal-Trend Decomposition Procedure Based on Loess. *Journal of Official Statistics*.
- Breiman, L. (2001). Random Forests. *Machine Learning*.
- Chen, Z., et al. (2022). Monitoring City Lights in India: A VIIRS-DNB-Based Assessment of Electrification and Economic Growth. *Sustainabil- ity*.
- Kulkarni, S., & Shinde, S. (2023). VIIRS Nighttime Light and Open-StreetMap for Urban Planning in Indian Mega-Cities. *ISPRS International Journal of Geo-Information*.
- Mundra, A., & Gupta, M. (2021). Geospatial Analysis of COVID-19 Impact on Urban Consumption using Nighttime Lights in Pune, India. *Geocarto International*.
- Sun, X., et al. (2019). Comparison of DMSP-OLS and VIIRS Nighttime Light Data in Modeling Urban Economic Activity in China. *Remote Sensing*.
- Wang, T., & Li, Q. (2020). Using VIIRS Nighttime Light Data to Evaluate the Urban Development Level of Coastal Cities. *IEEE Access*.
- Elvidge, C. D., et al. (2017). VIIRS Nighttime Lights: The Dawn of a New Era. *Remote Sensing*.
- Ministry of Housing and Urban Affairs (MoHUA), "Smart Cities Mission: Strategy and Operational Guidelines," Government of India, 2024.
- Chen, J., et al. (2022). Anomaly Detection in VIIRS Nighttime Light Time Series Using a Hybrid Deep Learning Model. *International Journal of Remote Sensing*.
- T. Chen and C. Guestrin, "XGBoost: A Scalable Tree Boosting System," in *Proc. 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2016, pp. 785–794.
- United Nations, "SDG 11: Make cities and human settlements inclusive, safe, resilient and sustainable," *Sustainable Development Goals Progress Report*, 2023.
- N. Gorelick et al., "Google Earth Engine: Planetary-scale geospatial analysis for everyone," *Remote Sensing of Environment*, vol. 202, pp. 18–27, 2017.