TECHNICAL NOTE Nº IDB-TN-03319

# Putting Numbers on a Continent: Spatial Measurement of Economic Activity in Amazonia

Pablo M. García
Roberto Duran-Fernandez
David Figueroa

Inter-American Development Bank
Productivity, Trade, and Innovation Sector

March 2026

IDB

# Putting Numbers on a Continent: Spatial Measurement of Economic Activity in Amazonia

Pablo M. García
Roberto Duran-Fernandez
David Figueroa

Inter-American Development Bank
Productivity, Trade, and Innovation Sector

March 2026

IDB

# Putting Numbers on a Continent: Spatial Measurement of Economic Activity in Amazonia

Pablo García[1], Roberto Duran-Fernandez[1] & David Figueroa[2] [3]

**Abstract**

Despite its continental scale and global relevance, Amazonia lacks spatially disaggregated, consistent GDP measures, leaving its economy largely unmeasured at the scale where development and environmental decisions are made. This paper introduces a scalable framework to estimate subnational economic activity by combining machine learning, nighttime lights (NTL), and spatial data within a unified 5×5 km grid. We develop Random Forest models integrating NTL, electricity consumption, population density, geographic features, spatial spillovers, and temporal dynamics. To assess the contribution of different information sources, we estimate six model specifications that progressively incorporate additional covariates, including a comparison between raw luminosity and satellite-derived GDP measures. To ensure realistic performance, we apply rolling-origin temporal cross-validation to prevent temporal leakage. Our findings show that combining NTL with spatial structure and temporal dynamics significantly improves subnational GDP estimation relative to luminosity-only methods. Crucially, the paper distinguishes forecasting over time from generalizing across space: while single-country models outperform in temporal prediction, multi-country training markedly improves spatial generalization in data-scarce regions. These findings highlight the importance of distinguishing temporal forecasting from spatial generalization when applying machine-learning models to economic measurement in large and heterogeneous territories.

JEL Codes: C53; C55; Q56; E01

Keywords: *Amazonia, Nighttime Lights (NTL); Sustainable Development*

[1] Inter-American Development Bank
[2] Université Paris 1 Panthéon Sorbonne
[3] This research document reflects the views of its authors and does not represent the official positions, policies, or opinions of the institutions with which they are associated.

## 1. Introduction

If Amazonia were a country, it would be the sixth-largest territory on Earth, an area larger than Australia and the entire contiguous United States. Although definitions vary across hydrological, ecological, and socio-environmental criteria, Amazonia broadly refers to the vast tropical forest and river basin system spanning nine countries in South America. It covers 7 to 8.5 million square kilometers, the region combines continental scale with exceptional ecological, cultural, and climatic significance. Yet, despite its magnitude and global importance, remarkably little is known about its economy.

Amazonia is home to millions of people, vibrant local economies, and long-standing trade networks, but in statistical terms it remains one of the world's largest economic blind spots. Official datasets typically report GDP only at highly aggregated administrative levels, masking the local dynamics that shape livelihoods, connectivity, and opportunity. Similar limitations of subnational economic data have been documented in other data-scarce and spatially heterogeneous contexts, where policymakers operate with coarse and incomplete information despite governing large territories (Henderson et al., 2012; Gibson et al., 2021).

This information gap has tangible consequences. A region of continental scale, accelerating development pressures, and global environmental relevance is managed using data that provide, at best, a coarse approximation of economic reality. Infrastructure investments are planned with limited spatial visibility, environmental policies rely on aggregates too broad to guide local interventions, and development strategies often depend more on anecdotal evidence than systematic measurement. (Donaldson and Storeygard, 2016).

The central question of this paper emerges directly from this challenge: How can economic activity be measured in vast territories where traditional statistical systems provide only coarse spatial coverage? Amazonia provides a particularly stark example of this problem. Its development trajectory, infrastructure priorities, conservation strategies, and social service provision all depend on understanding where people live and where economic activity occurs. Yet official GDP statistics are typically reported at the level of states, departments, or regions, but these units are often larger than many countries. As a result, they obscure profound economic heterogeneity. In Brazil, for example, the State of Amazonas is roughly the size of Colombia, yet two-thirds of its population lives in Manaus. State-level statistics therefore describe Manaus, not the interior. Similar patterns are observed across developing regions characterized by large territories and uneven settlement patterns (Bickenbach et al., 2016; Gibson et al., 2021).

Estimating the economic output of Amazonia presents significant methodological challenges (Durán-Fernandez, R., & de Carvalho Coutinho, 2025). This is compounded by the structural constraints faced by conventional data collection systems. Household surveys and administrative records are costly and difficult to implement in remote areas characterized by low population density, long distances, fragmented infrastructure, and uneven institutional capacity. As a result, the regions facing the greatest development, connectivity, and environmental pressures are often those that are least measured (Jean et al., 2016).

Against this backdrop, non-traditional data sources offer a promising and scalable alternative. Advances in remote sensing, geospatial data, and machine-learning methods have made it possible to observe human presence and economic intensity using satellite-based NTL, electricity consumption proxies, gridded population datasets, and geographic information systems (GIS). Among these, NTL has been widely used as proxy for economic activity, capturing artificial illumination observable from space (Mellander et al., 2015). At the same time, luminosity alone performs poorly in rural, agricultural, and forested regions and

must be combined with additional contextual information to recover economic activity in heterogeneous landscapes (Bundervoet, Maiyo, and Sanghi, 2015; Bickenbach et al., 2016).

This paper builds on these advances to develop a scalable, data-driven framework for estimating subnational economic activity in Amazonia. We construct a high-resolution 5 × 5 km grid-based dataset, and integrate satellite-derived indicators, spatial spillovers, and geostatistical relationships using Random Forest models. To ensure realistic predictive performance, we implement a temporal cross-validation strategy that avoids unintended temporal leakage. The resulting framework generates spatially explicit and temporally consistent GDP estimates that can inform evidence-based policymaking and support sustainable development strategies across the region.

The paper contributes to the literature on spatial economic measurement in four ways. First, it develops a scalable framework for estimating subnational economic activity in data-scarce environments by integrating NTL with spatial, demographic, and geographic covariates within a machine-learning architecture. Second, it generates an unprecedented high-resolution dataset of economic activity for Amazonia on a 5 × 5 km grid. Third, it introduces a downscaling procedure that allocates predicted subnational GDP to grid cells using luminosity-based fractional weights, producing a harmonized spatial representation of economic activity. Fourth, the analysis shows that multi-country training improves spatial generalization in large heterogeneous territories. This result highlights a key methodological contribution of the paper: forecasting over time and generalizing across space are fundamentally different predictive tasks. Countries with long GDP histories provide richer temporal information but require models to predict across time under relatively stable structural conditions. In contrast, multi-country models expose the algorithm to a wider range of spatial heterogeneity, improving their ability to generalize to data-scarce regions. Distinguishing between these two prediction tasks is therefore critical for the design and evaluation of machine-learning models applied to subnational economic measurement.

Naturally, several limitations remain. The availability and consistency of subnational GDP data differ across countries, constraining comparability and influencing model generalization. NTL-based proxies may underrepresent production occurring without artificial light, such as informal, agricultural, or forest-based activities, and may overstate sectors with high luminosity but limited value added. Population datasets also tend to undercount residents in remote rural areas (Láng-Ritter et al., 2025), introducing potential bias. In addition, the conversion of official GDP statistics into constant U.S. dollars requires assumptions that may not fully capture subnational price variation.

Despite these caveats, the methodological approach presented here offers a replicable and scalable framework for generating consistent, spatially detailed economic estimates in data-scarce environments. While developed for Amazonia, the approach is broadly applicable to other large territories where official statistics remain sparse or uneven, providing a foundation for improved measurement of economic activity across space.

## 2. Literature Review

The use of NTL as a proxy for economic activity has gained increasing attention in the development, regional economics, and economic geography literature. Foundational contributions by Henderson, et al., (2012) and Mellander et al. (2015) established the empirical framework linking satellite-based measures of artificial illumination to economic output. Their approach, which exploits the spatial continuity of luminosity as an observable proxy for human activity, underpins much of the subsequent literature, including the harmonized global gridded datasets developed by Chen et al. (2021a, 2022). Earlier, Chen and Nordhaus (2011) demonstrated that NTL could substitute for conventional economic statistics in data-scarce countries, highlighting their potential for cross-country measurement. Donaldson and Storeygard (2016) provide a broader survey of the satellite data revolution in economics, Gibson et al., 2021 offer a comprehensive review of NTL data sources, calibration challenges, and empirical applications. At the same time, a growing body of evidence emphasizes that NTL should not be interpreted as a standalone measure of economic output.

A central limitation of NTL-based approaches is their uneven performance across space. Luminosity captures urbanization, industrial activity, and electrification particularly well, but it performs poorly in rural and low-density areas dominated by agriculture, forestry, or informal activities. Using detailed subnational data from Indonesia, Gibson et al. (2021) show that NTL-based GDP estimates systematically underperform in rural regions, leading to spatially correlated measurement error. Importantly, this limitation reflects structural features of economic activity rather than sensor noise or data quality issues.

Evidence from other contexts reinforces this conclusion. Bickenbach et al. (2016) document substantial cross-country heterogeneity in the relationship between NTL and GDP, showing that correlations differ markedly between urban and rural areas in both India and Brazil. While NTL-based estimates perform well in cities, they become less reliable outside urban cores. These findings suggest that the mapping between luminosity and economic activity is highly context-dependent and sensitive to settlement patterns and sectoral composition.

Although few studies focus explicitly on Amazonia, available evidence points in a similar direction. Using subnational data from Colombia, Pérez-Sindín et al. (2021) find that NTL-based GDP estimates perform worst in sparsely populated, forest-covered local jurisdictions, particularly in the Colombian Amazon. The reduced performance is attributed to a combination of dense forest canopy, low population density, and limited electrification, all of which weaken the observable relationship between artificial light emissions and economic output (da Rocha Bragion et al., 2019).

In response to these limitations, literature emphasizes the value of integrating additional covariates alongside NTL to improve estimation accuracy. Variables capturing population distribution, land use, agricultural activity, and vegetation cover help recover economic dynamics that generate little artificial light and reduce systematic spatial bias (Bundervoet, Maiyo, and Sanghi, 2015; Bickenbach et al., 2016; Gibson et al., 2021). The underlying rationale is that different land-cover types and settlement structures are associated with distinct production technologies and economic intensities, which cannot be inferred from luminosity alone.

Building on this insight, several studies incorporate geographic and infrastructure variables into NTL-based estimation frameworks. Jin et al. (2023), for example, show that combining NTL with elevation, road networks, and accessibility measures significantly improves subnational GDP estimates in Thailand,

enhancing performance in both urban and rural areas. These results highlight the importance of accounting for physical geography and connectivity when interpreting satellite-based proxies of economic activity and support modeling approaches that integrate NTL within a broader spatial and infrastructural context.

Taken together, this literature establishes three empirical regularities that directly inform our modeling strategy. First, the elasticity between NTL and GDP is not constant across contexts. It varies systematically with income levels, urbanization patterns, sensor technology, country size, and administrative scale (Henderson, et al, 2012; Galimberti, 2020; Bickenbach et al., 2016), and may even reflect political distortions in certain institutional settings (Martínez, 2022). This heterogeneity cautions against simple pooled linear calibrations and motivates the use of flexible, nonlinear estimators trained on country-level data. Second, sensor technology materially affects predictive performance. VIIRS-based measures substantially outperform earlier DMSP series at subnational scales, while DMSP often exhibits weak or insignificant relationships at lower administrative levels (Gibson, et al, 2021). [4] The use of the harmonized Chen et al. (2021a) series is therefore motivated by the need for a temporally consistent, sensor-calibrated panel covering 2000–2019. Third, machine-learning approaches consistently outperform NTL-only regressions when richer covariate sets are available. Applications ranging from transfer learning (Jean et al., 2016) and neural networks (Yeh et al., 2020; Khachiyan et al., 2022) to Random Forest methods (Rossi-Hansberg and Zhang, 2025) demonstrate substantial predictive gains from incorporating demographic, environmental, and infrastructural information. These findings collectively support the multi-covariate Random Forest framework adopted in this paper.

Among these contributions, Rossi-Hansberg and Zhang (2025) provide the closest methodological benchmark. They train a Random Forest model using NTL, population, $CO_2$ emissions, and land-use covariates to generate gridded GDP estimates at the global scale. Our paper complements their global framework by focusing on a spatially heterogeneous region where the central challenge is not global coverage but recovering fine-scale economic variation within a vast territory. We apply this approach to Amazonia at a 5 × 5 km resolution -between five and twenty times finer than their grid- and incorporate region-specific ecological and institutional covariates, including deforestation, illegal mining, protected areas, and hydroelectric infrastructure. In addition, we explicitly model spatial spillover variables that are central to the economic geography of Amazonia.

The design of our explanatory variables is conceptually informed by the New Economic Geography (NEG) literature (Fujita, Krugman, and Venables, 1999). Population density proxies agglomeration forces; road and river accessibility capture market-access gradients; mining and oil-extraction indicators reflect patterns of factor-endowment specialization; and spatial lags of NTL and population approximate upstream–downstream economic linkages across the river system. While these variables are motivated by economic geography theory, we do not estimate structural parameters of a formal NEG model. Rather, the Random Forest is trained to predict subnational GDP out of sample. The selection of covariates reflects substantive economic intuition about spatial development patterns, without imposing a structural data-generating process.

[4] DMSP is an older U.S. Air Force satellite sensor that captured nighttime lights at coarse resolution without radiometric calibration, providing long historical coverage but with limitations in precision. VIIRS is a newer, fully calibrated NOAA/NASA sensor with much higher spatial resolution and a wider dynamic range, producing cleaner and more reliable nighttime light measurements.

## 3. Study Area and Data Sources

### *3.1. Definitions of the Amazonia Region*

Defining the Amazonia region is inherently complex, as alternative delineations exist depending on physical geography, biogeographical boundaries, and socio-environmental criteria. These variations result in estimates of Amazonia's area ranging from 7.2 million to 8.4 million km². The literature identifies at least three main approaches for delineating Amazonia (González Saldarriaga, et al., 2025) (Figure 1):

a) ***Amazon River Basin.*** This definition encompasses the entire hydrological basin that drains into the Amazon River, extending between the Guiana Shield and central Brazil (north-south) and from the Atlantic Ocean to the Andes (east-west). The estimated area of this basin is approximately 6.7 million square kilometers (WWF, 2023).

b) ***Biogeographic Limits*****:** This approach defines Amazonia based on its dominant tropical rainforest ecosystem, including dense humid forests, seasonally flooded forests, and patches of savanna, bamboo, and palm forests. Estimates suggest an area between 6.7 and 7.2 million km² (WWF, 2023).

c) ***RAISG.*** The Amazonian Network of Socio-Environmental Information (RAISG) was established in 2007 to provide an integrated and decentralized knowledge base for Amazonia. It defines Amazonia as an area covering 8.4 million km², making it the largest of the three definitions. This definition is widely used in geo-statistical, academic, and official studies, including this paper (RAISG, 2020).

The Amazonian region spans across nine national jurisdictions, including eight sovereign Latin American countries (Bolivia, Brazil, Colombia, Ecuador, Guyana, Peru, Suriname, and Venezuela), as well as the overseas department of French Guiana (France).[5] Regardless of the definition employed, Amazonia is comparable in size to the continental United States or Australia and would rank as the sixth-largest territory in the world if it were a sovereign entity.

### *3.2. Study Area and Grid Construction*

This paper follows the RAISG definition as the primary geographical reference. However, the study area in this analysis extends beyond its strict boundaries. To ensure consistency in economic estimation, the study incorporates entire subnational administrative units that intersect with the RAISG-defined Amazonia, even if portions of these regions fall outside its officially designated limits. This approach allows for a more accurate comparison with official subnational economic statistics while maintaining alignment with established geographic definitions. Specifically, the study incorporated complete local level subnational jurisdictions that had at least 2% of their area overlapping with the RAISG boundary. This threshold was determined to account for polygon slivers that might be unintentionally adding municipalities into the RAISG boundary (Delafontaine et al., 2009). This adjustment was made to prevent artificial truncation of economic data, particularly Gross Domestic Product (GDP) estimates, which are typically reported at the municipal level (Figure 2).

[5] RAISG do not include the French Guina in its primary study.

The analytical framework was structured using a 5 × 5 km grid, designed to ensure spatial consistency across national boundaries. The grid construction was carried out in QGIS using the SIRGAS 2000 projection system, which was selected due to the extensive geographic coverage of the study area.

Each grid cell was systematically assigned official administrative codes, following a classification system inspired by the European Union *Nomenclature of Territorial Units for Statistics* (NUTS) (European Union; 2024).

a) **NUTS0** represents the national level.

b) **NUTS1** corresponds to the first subnational administrative level.

c) **NUTS2** accounts for the second level subnational jurisdiction in Bolivia, Ecuador, and Peru (*i.e.*, regional administrative units according to Eurostat definitions).

d) **NUTS3** represents the second level subnational jurisdiction in Brazil, Colombia, Guyana, Suriname, and Venezuela, and the third level subnational jurisdiction in Bolivia, Ecuador, and Peru (*i.e.*, local level subnational jurisdictions according to Eurostat definitions).

The grid identifies NUTS0, NUTS1, and NUTS3 for all cells. This approach harmonizes the spatial framework for GDP estimation across the Amazonia region, facilitating robust comparisons across national jurisdictions while preserving the highest possible level of spatial disaggregation. Table 1 presents the definition for each NUTS in each country of the study area.

It is important to clarify that subnational socioeconomic data is typically produced at the second level subnational jurisdiction (NUTS1); however, data availability for lower levels (NUTS2, NUTS3) is inconsistent across countries. Moreover, for countries with two or more subnational layers, it is rare to find a complete geostatistical framework encompassing all levels.

### *3.3. GIS Data Sources*

This study integrates multiple GIS datasets to estimate economic activity in the Amazonia region, combining GDP, population, deforestation, and various geographic variables. The primary economic indicator is the gridded GDP dataset from Chen et al. (2021a), which provides annual estimates from 1990 to 2019 at a 1 km resolution, calibrated by harmonizing DMSP and VIIRS NTL data.

Population data are obtained from the Global Human Settlement Population (GHS-POP) dataset (Carioli et al., 2023), which provides quinquennial estimates from 2000 to 2020 at a spatial resolution of 100 m × 100 m. To maintain temporal consistency, missing intermediate years are estimated using geometric interpolation, which assumes constant annual growth rates between GHS-POP anchor years (2000, 2005, 2010, 2015, and 2020). This assumption is standard in the absence of annual high-resolution population data for the region, but it may be violated in contexts where demographic dynamics are shaped by abrupt infrastructure expansion, conservation-related displacement, migration linked to extractive activities, or illegal-economy shocks. To partially mitigate this limitation, the empirical specification incorporates temporal lags and two-year moving averages of population and density variables, allowing the model to capture short-run persistence and smooth transitory fluctuations. Future research may improve upon this constraint by integrating alternative demographic proxies, such as mobile-phone activity data or satellite-derived annual settlement footprints.

Additional datasets capture infrastructure, land use, and environmental conditions. Annual deforestation data from RAISG (2024a) provides annual observations from 2000 to 2019, reflecting land use changes that may influence regional economic patterns. Road networks, obtained from RAISG (2023c) and Rocha et al. (2023), offer insights into connectivity and accessibility. NTL data from Chen et al. (2020, 2021b) capture spatial variation in artificial illumination and economic intensity.

Land-use variables include urban extent data from Zhao et al. (2022) and harvested area from Potapov et al. (2021). Elevation data obtained from WWF (2006) helps assess topographical constraints on development. The dataset also integrates indicators of extractive industries and environmental protection, including legal and illegal mining, hydroelectric plants, protected areas, and oil extraction zones compiled from RAISG (2023a; 2023b; 2024b; 2024c; 2024d).

These diverse datasets collectively provide a comprehensive foundation for estimating GDP at a granular level, leveraging both traditional economic indicators and spatial proxies for economic activity. A summary of geographic data is presented in Table 2.

### *3.4. Official GDP Data and Standardization*

The availability of official subnational GDP data varies across Amazonian countries. Brazil and Peru provide the most comprehensive coverage, with GDP data available at the local administrative level from 2002–2019 and 1993–2018, respectively. In Brazil, GDP statistics are published by the Brazilian Institute of Geography and Statistics (IBGE) and IJSN (2024), are available at both the NUTS1 and NUTS3 levels (*municipios*). In Peru, subnational GDP estimates are available at the NUTS1 and NUTS2 levels (*distritos*), with figures estimated by the IDB.

Colombia publishes subnational GDP data at NUTS1 and NUTS3 levels (*municipios*) for the period 2011–2019. Ecuador had GDP estimates for 2018 and 2019 at NUTS2 level (*canton*), but these were excluded from the models, as several methodological approaches used in the study require historical data to compute moving averages. Including only two recent years would have introduced inconsistencies in the time-series and reduced the reliability of the estimations. Meanwhile, Guyana, Venezuela, and Suriname do not have any publicly available subnational GDP data.

A key challenge in integrating official GDP figures was the heterogeneity of currencies, as all values were originally reported in current local currency units (LCU). This posed difficulties for direct comparison and for training machine learning models since the results would not be interpretable in a consistent monetary unit. To maintain consistency across currencies and sources, all GDP figures in the study were standardized to real 2017 USD. Since Chen et al.'s (2021a) GDP was already expressed in real 2017 USD, only the official GDP required adjustment (See Appendix A).

### *3.5. Constructed Data*

Several additional variables were constructed by interacting with the original geographic datasets to enhance the models' predictive power and capture more nuanced spatial-economic relationships. As detailed in Table 3, these variables were created by combining multiple data sources to reflect relative rather than absolute values. For instance, population density was calculated by dividing the total population by the area of each grid cell, providing a more accurate representation of population distribution. Similarly,

road density was determined by summing the total length of roads within a cell and dividing it by its area, allowing for a standardized measure of transportation infrastructure across different regions.

These interactions extended beyond basic density metrics, incorporating relationships between variables such as urbanization levels, accessibility, and environmental constraints. This approach enabled the models to better account for the complex spatial heterogeneity that characterizes economic activity in Amazonia, improving their ability to capture local economic dynamics in a geographically diverse region.

### *3.6. Data Integration into the Grid*

Most of the datasets used in this study are derived from raster data, while the analytical framework is based on a 5 × 5 km vector grid. To harmonize these different data structures, a consistent methodology is required to integrate raster-based values into the grid while preserving the spatial and economic significance of each variable. Moreover, even vector data does not always align perfectly with grid cells, for example, roads may only partially intersect a cell rather than fully covering it. As a result, a structured approach is necessary to ensure that both raster and vector data are accurately represented within the grid, maintaining spatial consistency.[6]

For GDP and population, a summation method is applied, aggregating values from the original raster datasets within each grid to accurately represent economic activity and demographic distribution. Since population data is only available at quinquennial intervals, missing values for intermediate years are estimated using geometric interpolation, ensuring temporal continuity.

In addition to GDP and population, other geographic and economic variables are integrated into the 5 × 5 km grid using tailored aggregation techniques. For roads, the total length of segments within each cell is summed to reflect infrastructure density, even when road networks only partially intersect a cell. NTL intensity and electricity consumption values are also aggregated by adding all pixel values within each cell, providing a representative measure of economic activity. For land use variables, such as urban area and harvested area, values are averaged across pixels in each grid cell to capture land coverage patterns. Elevation is summarized by using three metrics (minimum, maximum, and mean elevation) to account for topographic variation.

For categorical spatial features, binary presence-absence indicators are generated to identify whether a given grid cell contains specific land uses or activities. Features such as illegal mining, legal mining, protected natural areas, hydroelectric plants, and oil extraction zones are each assigned to a binary variable indicating their presence. Additionally, a count variable is used to record the total number of such features per cell. To account for influences beyond cell boundaries, an extra indicator is included to reflect whether a hydroelectric plant is located within a 50 km radius of the cell. These data integration techniques ensure that both continuous and categorical variables retain their spatial and economic relevance within the grid-based framework, supporting a consistent and comparable approach across datasets.

This methodology is also applied to NUTS3 subnational units. Instead of aggregating values within fixed grid cells, the data are aligned with administrative boundaries, and appropriate summarization techniques are applied. For continuous variables such as GDP, population, and NTL, aggregation involves summing

[6] Raster data represent information as a grid of pixels, where each cell stores a value (*e.g.*, light intensity, land use). Vector data, by contrast, use points, lines, and polygons to depict discrete features such as cities, roads, or administrative boundaries.

up raster pixel values within each administrative unit to capture broader economic and demographic patterns. For infrastructure data, such as roads and electricity networks, the total length or density within each unit is computed to reflect connectivity. Similarly, land use and environmental features are averaged across the full area of each subnational entity to ensure spatial consistency. Categorical features, such as the presence of protected areas or extractive industries, are represented through binary indicators or total counts per unit. Proximity-based indicators, such as the presence of hydroelectric plants within a specified distance, are also included to reflect extended spatial influences.

## 4. Methodology

### *4.1. Machine Learning Approach*

The primary objective of the analysis is to produce GDP estimates at the lowest feasible subnational level while ensuring consistency with available official statistics. Although NTL provide a useful proxy for the spatial distribution of economic activity, they do not perfectly recover absolute GDP levels. We therefore calibrate a predictive model using countries with official subnational GDP data and apply the trained model to jurisdictions where such data are unavailable.

To operationalize this strategy, we employ a Random Forest machine-learning framework that integrates spatial, demographic, environmental, and economic predictors described in Section 3. The dependent variable is official subnational GDP, observed for Brazil, Peru, and Colombia at the NUTS3-equivalent level (*municipios* in Brazil and Colombia; *distritos* in Peru; see Table 1). The dataset includes only jurisdictions within the designated study area and spans multiple years, as described in Section 3.2

Six model specifications are estimated, each incorporating alternative combinations of covariates and validation procedures to assess predictive robustness. Estimation relies exclusively on jurisdictions with observed subnational GDP. The trained models are then used to generate GDP predictions for countries lacking official subnational data. Estimation and forecasting are conducted at the NUTS3-equivalent administrative level for all countries, except Bolivia and Ecuador, where NUTS2-equivalent units are used.

Predicted subnational GDP is subsequently distributed across 5 × 5 km grid cells using a mass-preserving allocation rule. Specifically, each unit's predicted GDP is downscaled in proportion to the within-unit spatial distribution captured by satellite-derived gridded GDP (Chen et al., 2021a), ensuring that the sum of grid-cell values equals the predicted total for each jurisdiction.

Although Random Forests do not estimate a parametric model with explicit coefficients, the underlying functional relationship can be represented in pseudo-parametric form to clarify the structure of the predictor groups and their interaction.[7]

**Equation 1.** Random Forest Structure

$$y_{i,t} = f\left(PAC_{i,t}, STR_{i,t}, STR^{D}_{i,t}, STR^{S}_{i,t}, NTA_{i,t}, NTA^{S}_{i,t} ECO_{i,t}, GEO_{i,t}, INT_{i,t},\right)$$

[7] The Random Forest algorithm builds an ensemble of decision trees using random subsamples of both data and predictors; each tree partitions the feature space into homogeneous regions, and the final prediction is obtained by averaging across all trees to minimize overfitting, out-of-sample predictive accuracy, and capture complex, non-linear interactions among variables.

Where $\hat{y}_{i,t}$ denotes the predicted GDP of spatial unit *i* in year *t*, and $f(\cdot)$ is a non-parametric function learned by the Random Forest algorithm. The explanatory variable groups are defined as follows:

- ***PAC* — Primary Alignment Controls.** This set includes NTL-based GDP from Chen et al. (2021a), the proportion of NUTS3 GDP relative to NUTS1 GDP, and country dummies.
- ***STR* — Structural Baseline Variables**. These include population, population density, and deforestation, which serve as baseline indicators capturing fundamental structural conditions of local economic activity.
- ***STR*$^{D}$ — Structural Dynamics**. These variables capture the temporal dynamics of *STR*, including lags, growth rates, and moving averages that reflect persistence and change over time.
- ***NTA* — Nighttime Activity Indicators.** This group includes NTL and electricity consumption, both of which serve as high-resolution proxies for local economic activity (Chen et al. 2021a).
- ***ECO* — Economic Level & Trend Indicators.** This set comprises economic indicators derived from Chen et al. (2021a), including levels, lags, growth rates, and density measures that capture macro-to-local economic variation.
- ***GEO* — Geographic & Land Conditions.** These variables describe the physical and territorial context of economic activity, including topography, land use, infrastructure, and accessibility.
- ***INT* — Interactions.** These include cross-variable and country-specific interactions that capture structural heterogeneity and non-linear joint effects.

For both *STR* and *NTA*, we also incorporate nine spatial spillover metrics that summarize the influence of neighboring units, which is maximum, minimum, and mean values and their corresponding differences relative to the focal unit (***STR*$^{S}$** and ***NTA*$^{S}$**, respectively). The inclusion of these metrics follows the spatial-dependence literature (Sekulic et al., 2020).

We estimate six versions of Equation 1, each incorporating different combinations of these data sources, to evaluate the contribution of each variable group to model performance. This comparative approach enables a systematic assessment of how different inputs affect GDP estimation accuracy, guiding the selection of the most reliable methodology. Table 4 presents an exhaustive description of the variable groups defined above, the specific variables included within each group, and the model versions in which they appear. The performance of all models is summarized in Table 5.

All performance statistics in Table 5 are computed at the NUTS3 level, which is the geographic unit at which official GDP data are available and at which the Random Forest models are trained and validated. The subsequent downscaling of predicted NUTS3 GDP to the 5 × 5 km grid is implemented only after model evaluation. This sequencing ensures that the Chen et al. (2021a) estimates used as spatial allocation weights do not influence or contaminate the reported predictive accuracy measures.

*Model 1: Baseline.*

**Equation 2.** Random Forest Structure in Model 1

$$y_{i,t} = f(PAC, STR_i)$$

The model includes the structural baseline variables (*STR*) along with the primary alignment variables (*PAC*), which serve as controls analogous to those in the fixed-effects specification. A standard 70/30 random split is used to partition the data into training and testing sets. This baseline model establishes a reference point for assessing how the incorporation of additional data sources and methodological refinements enhances predictive performance in the subsequent models.

*Model 2: Baseline + Temporal Cross Validation.*

**Equation 3.** Random Forest Structure in Model 2

$$y_{i,t} = f\left(PAC_{i,t}, STR_{i,t}\right)$$

This builds upon Model 1 by incorporating a temporal validation approach to account for the panel structure of the data. While it uses the same predictor variables as Model 1, it explicitly addresses the presence of temporal autocorrelation in GDP estimates. Following Roberts et al. (2017), this model applies Rolling Origin temporal cross-validation to measure the impact of temporal dependence on the data. This approach enhances the robustness of the Random Forest predictions by ensuring that the model generalizes well over time, rather than simply capturing patterns from a random training-test split. Error metrics are expected to be higher than those in Model 1, but they offer a more reliable indication of predictive accuracy. This trade-off ensures that the model's performance is evaluated in a way that better reflects real-world temporal dynamics.

*Model 3: Geographic & Land Conditions Variables*

**Equation 4.** Random Forest Structure in Model 3

$$y_{i,t} = f\left(STR_{i,t}, STR^{D}_{i,t}, NTA_{i,t}, GEO_{i,t}\right)$$

Model 3 excludes the GDP estimates from Chen et al. (2021a), relying solely on nighttime activity indicators (*NTA*), which include NTL and electricity consumption, along with the full set of variables that capture the human and physical geographic characteristics of the territory (*GEO*). This specification follows the methodology proposed by Jin et al. (2023). To ensure robustness against temporal autocorrelation, the model again incorporates Rolling Origin temporal cross-validation. By removing NTL-based GDP estimates and focusing exclusively on geographic determinants, this model evaluates the extent to which spatial features alone can account for variation in economic activity, providing an alternative benchmark against which the performance of the other models can be compared.[8]

*Model 4: Geographic & Land Conditions Variables + Neighboring Effects*

**Equation 5.** Random Forest Structure in Model 4

$$y_{i,t} = f\left(STR_{i,t}, STR^{D}_{i,t}, STR^{S}_{i,t} NTA_{i,t} NTA^{S}_{i,t}, GEO_{i,t}, INT_{i,t}\right)$$

[8] As part of the spatial estimation process, Area-to-Area Kriging was initially tested as an alternative interpolation method for GDP estimation following Yin et al. (2023). However, this approach produced negative GDP estimates in some cases, due to the substantial economic disparities across different countries within the Amazonia region. Given these inconsistencies, Model 3 does not use Kriging and relies exclusively on the GDP estimates generated by the Random Forest model using geographic variables.

Model 4 extends the geographic variables model (Model 3) by incorporating neighboring effects for both the structural variables ($STR^S$) and the nighttime activity indicators ($NTA^S$). It also adds additional controls, including the temporal dynamics of structural variables ($STR^D$) and interaction effects (*INT*) across selected structural, nighttime activity, and geographic variables. The inclusion of neighboring effects accounts for the influence of adjacent jurisdictions, enabling the model to capture spatial dependence. This structure allows the Random Forest algorithm to learn complex spatial-dependence patterns and non-linear relationships, acknowledging that economic activity is inherently spatially correlated and structurally non-linear. As in Model 3, the GDP estimates from Chen et al. (2021a) are excluded to isolate the predictive contribution of GIS-based variables alone. Rolling Origin temporal cross-validation is again applied to ensure robustness to temporal autocorrelation.

*Model 5: Adjusted Chen's GDP*

**Equation 6.** Random Forest Structure in Model 5

$$y_{i,t} = f\left(PAC_{i,t}, STR_{i,t}, STR^D_{i,t}, STR^S_{i,t}, ECO_{i,t}, GEO_{i,t}, INT_{i,t},\right)$$

Model 5 serves as a benchmark for evaluating the predictive contribution of the GDP estimates from Chen et al. (2021a) relative to their underlying nighttime activity inputs. Whereas Models 3 and 4 excluded Chen's GDP metric to assess whether raw NTL, electricity consumption, and geographic a land conditions variables could independently predict official GDP, Model 5 reincorporates the Chen estimates directly as predictors. This specification allows the Random Forest to learn and correct systematic deviations between Chen's GDP and official statistics by combining the Chen-derived values with the panel-alignment controls (*PAC*) and the variables capturing the temporal dynamics of Chen's estimates (*ECO*). Model 5 provides a direct benchmark for assessing the marginal value of the Chen GDP estimate relative to its underlying NTA inputs; and allows to test whether Chen's metric can be enhanced through a machine-learning approach that integrates its temporal behavior with a richer set of contextual variables. As in the previous models, Rolling Origin temporal cross-validation is applied to ensure that predictive performance is evaluated under realistic temporal conditions.

*Model 6: Full Model.*

**Equation 7.** Random Forest Structure in Model 6

$$y_{i,t} = f\left(PAC_{i,t}, STR_{i,t}, STR^D_{i,t}, STR^S_{i,t}, NTA_{i,t}, NTA^S_{i,t} ECO_{i,t}, GEO_{i,t}, INT_{i,t},\right)$$

This model integrates all key components from previous models. By combining these elements, the model aims to maximize predictive accuracy while capturing both spatial and temporal dependencies in economic activity. As in the other models, Rolling Origin Temporal Cross-Validation. This model represents the most comprehensive approach, leveraging both spatial and economic predictors while accounting for spatial dependence and temporal dynamics.

*4.2. Additional Robustness Estimations*

In addition to the six baseline model specifications described above, we conduct four complementary robustness exercises to assess the stability of the predictive framework under alternative validation schemes.

First, we restrict the estimation sample to Brazil. As noted earlier, Brazil provides the most complete and internally consistent subnational GDP series in the region, with annual observations available for all *municípios* (NUTS3) over an extended period. This yields a balanced panel with full geographic coverage and substantial temporal depth. We therefore re-estimate all six model specifications using the Brazilian sample exclusively. The Brazil-only estimation serves two purposes: it isolates model performance within a single, internally coherent institutional context, and it provides a benchmark against which the costs and benefits of pooling across heterogeneous national systems can be evaluated. The results are reported in Table 5.

Second, we implement a Leave-One-Country-Out (LOCO) cross-validation procedure to evaluate the model's ability to generalize across national contexts. In each iteration, one country (Brazil, Colombia, or Peru) is excluded from the training sample, and Model 5 (the preferred consolidated specification) is estimated using the remaining two countries. The trained model is then evaluated against the excluded country's official subnational GDP data. This procedure directly tests out-of-country predictive performance, rather than temporal forecasting within a pooled sample. LOCO represents the most demanding generalization test in the validation hierarchy, as it requires the model to predict an entire national context without exposure to any training observations from that country. Performance degradation relative to the temporal cross-validation baseline therefore provides a direct measure of the structural specificity of the learned relationships. Predictive accuracy is computed annually and summarized using overall $R^2$ and RMSE across the evaluation period (Table 6).

Third, we conduct a subnational leave-one-out procedure at the NUTS1 level. In this exercise, one NUTS1 unit is excluded from the training sample at a time, and the model is estimated using the remaining subnational units across all countries. Predictions are then generated for the NUTS3 units from the excluded NUTS1 unit and evaluated against official GDP data. This procedure is repeated sequentially for each NUTS1 unit in the estimation sample. The subnational analysis complements the LOCO by testing within-country spatial generalization at a finer geographic scale. Unlike the country-level exercise, it retains observations from the held-out country while withholding only a single macro-region, thereby isolating the model's ability to infer subnational economic patterns from the broader territorial context rather than from country-level structural similarity. Potential spatial leakage, where spatial lag variables for units adjacent to the excluded NUTS1 might inadvertently incorporate information from the withheld area, is mitigated by computing spatial lags using only training-sample units in each iteration. As in the LOCO, predictive performance is evaluated annually and summarized using overall $R^2$ and standard error metrics (Table 7).

Finally, we examine whether model residuals exhibit systematic spatial dependence. This diagnostic is essential because the spatial lag variables in Models 4 and 5 are intended to internalize spatial dependence within the covariate structure. If residuals remain significantly autocorrelated after inclusion of these variables, reported performance metrics may overstate true out-of-sample fit by implicitly relying on spatial information that would not be available in a genuine forecasting context. Conversely, a systematic decline in residual spatial autocorrelation across specifications provides direct evidence that the spatial controls are functioning as intended. Moran's I statistics are computed for official GDP, predicted GDP, and model

residuals across Model 1–5 specifications, for both the full multi-country sample and the Brazil-only sample. Spatial dependence is evaluated using four alternative weight matrices: row-standardized (W), variance-stabilizing (S), binary contiguity (B), and K-nearest neighbors with K = 4 (KNN). Statistical significance is assessed via permutation-based Z-scores (Tables 8–10). Robustness across alternative weight matrices provides an additional safeguard against artefacts arising from the choice of spatial definition.

### *4.3. Downscaling GDP Estimates to the Cell Level*

To achieve a high-resolution estimation of GDP at the cell level, a downscaling approach is applied using the outputs of the Random Forest models estimated at the subnational level. Models 1 to 5 are trained and validated using official subnational GDP data, as described in the previous sections. The estimated models are then used to generate GDP forecasts for all countries in the study area, including those without observed subnational GDP data, at the country-specific administrative level reported in Table 1 (*i.e.,* NUTS3-equivalent units for all countries except Bolivia and Ecuador, where NUTS2-equivalent units are used).

To distribute subnational GDP estimates across the 5 × 5 km grid, the analysis leverages satellite-derived gridded GDP from Chen et al. (2021a) to compute within-unit cell shares. Predicted subnational GDP is then allocated to grid cells in proportion to these shares, ensuring mass preservation by construction. This procedure scales local-level GDP predictions from the Random Forest models to the grid while preserving the spatial heterogeneity captured by NTL data.

Finally, a coefficient of variation (CV) analysis is conducted using model-based, downscaled grid-level GDP estimates for all countries in the study area, irrespective of whether official subnational GDP data were available for model estimation, to assess the temporal persistence of spatial economic structure. The CV captures the relative dispersion of these GDP shares over time, providing a scale-independent measure of structural stability.

Larger and more diversified economies, such as Brazil, Bolivia, and Peru, exhibit low CV values, implying that the spatial allocation of economic activity is remarkably persistent across years. Conversely, higher CV values in smaller economies or in countries with discontinuous official GDP series signal greater volatility in the regional distribution of output. Figure 3 illustrates this contrast, showing that spatial economic patterns are more stable in larger economies than in smaller, more volatile ones.

## 5. Results and Discussion

### *5.1. Comparative Performance of Model Specifications*

Across all three samples, Model 1, which relies on a random training-test split, achieves the lowest error metrics. This is expected because the random split allows the model to exploit temporal leakage, whereby information from future periods can inadvertently inform predictions for earlier years. Once temporal dependence is properly accounted for in Model 2, both RMSE and median absolute error increase. This increase reflects the fact that economic activity is persistently autocorrelated over time, and that random splits artificially inflate predictive performance. In this sense, Model 1 provides an upper bound on accuracy, whereas Model 2 provides a more realistic measure of temporal predictive performance. This confirms the importance of temporal cross-validation for evaluating GDP prediction models.

Models 3 and 4 remove the Chen et al. (2021a) GDP estimates entirely and rely exclusively on nighttime activity proxies (NTA) together with geographic and land condition variables (GEO). These models show systematically higher prediction errors than Models 1 and 2. This confirms that geographic fundamentals alone, even when augmented with neighboring effects and interaction terms, cannot fully explain short-run variation in official GDP. Nonetheless, Model 4 consistently outperforms Model 3, which suggests that neighboring variables and interaction effects meaningfully improve predictive accuracy. This result is consistent with the spatially clustered nature of economic activity, particularly in dispersed and sparsely populated regions such as the Amazon.

Reintroducing Chen's satellite-derived GDP estimates in Model 5 significantly increases predictive accuracy across all samples. RMSE falls to levels close to or better than those in Model 1, despite the use of temporally robust validation. This indicates that Chen's GDP estimates contain systematic predictive signal, and that the Random Forest is effective at learning and correcting systematic deviations between Chen's gridded GDP and official subnational statistics.

Finally, Model 6, the full model, achieves the strongest performance overall. It improves upon Model 5 by combining the information contained in Chen's GDP with the spatial information and contextual variables used in Models 3 and 4. The gains are strongest for the median absolute error (MAE), indicating that the full model reduces large prediction errors and produces more stable estimates across heterogeneous local contexts, an important property in a region as structurally diverse as Amazonia. Results for this comparison are presented in columns 1 and 2 of Table 5. Together, these results show that integrating satellite-derived GDP, nighttime activity proxies, and spatial-interaction features yield the most robust predictions for the Amazonia region.

### *5.2. Temporal Dynamics in the Brazil-Only Estimation*

Restricting the analysis to Brazil, as shown in columns 3 and 4 of Table 5, leads to higher predictive error than in the full Amazonia sample, even though Brazil provides the most complete and internally consistent dataset. This result may appear counterintuitive, but it is consistent with how temporal cross-validation behaves in long panels. Because Brazil contains many more years of GDP data, the Rolling-Origin procedure must generate a much larger number of genuinely out-of-sample forecasts. Each additional forecasting step increases the likelihood of encountering structural changes, economic shocks, or shifts in the relationship between NTL and GDP. As a result, the evaluation for Brazil is methodologically stricter and the higher predictive error reflects a more demanding validation environment rather than poorer model

quality. In contrast, the multi-country sample contains shorter time series for several countries, which requires far fewer temporal forecasting steps. The lower error observed there reflects an easier predictive environment rather than a stronger model.

A second factor is that geographic and infrastructural heterogeneity is lower within Brazil than across the entire Amazonia region. This reduces the informational content of spatial features and limits the model's ability to exploit cross-country contrasts. The satellite-derived GDP estimates from Chen et al. (2021a) are most informative when cross-country variation in electrification, density, and infrastructure is present in the training data; restricting the sample to Brazil removes this variation and attenuates their explanatory contribution. These two mechanisms, a stricter validation environment and reduced feature variance, together explain why Brazil-only error metrics are higher despite the superior data quality of the Brazilian series.

Importantly, although absolute errors are higher in the Brazil-only sample, the relative ordering of model performance remains stable. Models 1 and 5 deliver the best accuracy under random and temporal validation, respectively, and Model 6 remains the most robust specification overall. This stability confirms that the performance hierarchy identified in Section 5.1 is not an artefact of sample composition.

This distinction has a direct practical implication. Users forecasting future GDP for a single country with available historical data should rely on country-specific models, which offer tighter fit within a known structural context. Multi-country models are preferable for regional extrapolations and predictions in data-scarce countries, where broader structural relationships, learned from cross-country variation, provide stronger external validity and more reliable transfer to unobserved contexts.

### *5.3. Cross-Country and Subnational Generalization*

The results reported in Sections 5.1 and 5.2 establish that Model 5 and Model 6 achieve strong temporal predictive performance. The exercises in this section address a distinct and more demanding question: how well does the model generalize when required to predict economic activity in geographic contexts it has not observed during training? Two complementary tests are reported, LOCO cross-validation, and subnational Leave-One-NUTS1-Out cross-validation, each probing a different dimension of spatial generalization.

#### 5.3.1. LOCO cross-validation.

Table 6 reports LOCO results for Model 5. The central finding is a pronounced asymmetry across held-out countries. When Brazil is excluded from the training sample, out-of-country $R^2$ falls to 0.146, with an RMSE of 7,264.66 million BRL, a decline of more than 80 percentage points relative to the temporal cross-validation benchmark of 0.957. When Colombia is excluded, $R^2$ equals 0.372 (RMSE: 2,025.95), and when Peru is excluded, $R^2$ equals 0.412 (RMSE: 175.11).

Three mechanisms explain the sharp deterioration when Brazil is excluded. First, sample imbalance, as Brazil accounts for most training observations (all *municípios* from 2002 to 2019) removing it reduces effective sample size. Second, structural dominance, the NTL-to-GDP mapping learned by the model reflects Brazil's economic geography, a continental-scale economy combining metropolitan agglomerations, interior extractive belts, and dense river-basin infrastructure. Third, limited covariate

support, several feature combinations characteristic of Colombian and Peruvian Amazonia, smaller administrative units, lower average luminosity, and distinct sectoral composition, lie near the boundary of the feature space spanned by Brazilian observations. These mechanisms are conceptually distinct. Sample imbalance may be mitigated by expanded data coverage; structural dominance and limited covariate overlap reflect deeper limits to cross-country transferability when economic structures differ. Notably, the same structural dominance that explains why the Brazil-excluded LOCO performs poorly also explains why Brazil-only temporal validation is stricter than the pooled sample: the model's learned relationships are heavily indexed to Brazilian economic geography in both cases.

The Peru result ($R^2 = 0.412$; RMSE = 175.11) is comparatively encouraging. Peru's districts are smaller and GDP levels lower in absolute terms; consequently, RMSE remains modest even where relative fit is weaker. The Colombia result ($R^2 = 0.372$) lies between the extremes, consistent with Colombia's intermediate sample size and partial structural similarity to Brazil.

These findings qualify the claim of cross-country generalization. The model generalizes with moderate accuracy to Colombia and Peru when trained without them but extrapolates poorly to Brazil when trained exclusively on the two smaller economies. Predictions for countries outside the estimation sample (Bolivia, Ecuador, Venezuela, Guyana, and Suriname) should therefore be interpreted as conditional on structural comparability with the training countries. For Amazonia lowland territories with agro-extractive profiles, such as much of Bolivia and Ecuador, this comparability is plausible. For Venezuela and the Guianas, where macroeconomic dynamics and institutional structures diverge more substantially from the training sample, the extrapolation carries greater uncertainty. This hierarchy of confidence is reflected in the interpretation of the 5 × 5 km gridded estimates.

5.3.2. Subnational cross-validation.

Table 7 reports average RMSE and $R^2$ for each NUTS1 unit held out sequentially. The results reveal substantial regional heterogeneity and two categories of structural outliers.

**Outlier 1: Distrito Federal (Brazil).** The Distrito Federal generates the largest average RMSE in the sample (168,082 million BRL).[9] Two factors operate simultaneously. First, as Brazil's federal capital, the Distrito Federal concentrates administrative and public-sector value added, producing an unusually high GDP-to-luminosity ratio. A model calibrated primarily on Amazonian municipalities, where GDP is more linked to non-urban activity, may be underestimating this service-intensive economy. Second, the Distrito Federal is a scale outlier, its GDP is an order of magnitude larger than the next-largest NUTS1 unit in the sample, mechanically amplifying absolute prediction errors. Together, these factors make the Distrito Federal a structural boundary case for any framework relying on satellite-observable physical proxies. Its inclusion in performance summaries would distort aggregate error metrics without reflecting any generalizable weakness in the model's predictive logic.[10]

[9] Distrito Federal is composed by only 3 municipios so $R^2$ cannot be estimated.

[10] The Distrito Federal is located at the southeastern boundary of the study area and, together with La Paz, represents one of only two jurisdictions facing comparable structural conditions: large administrative centers whose economic profiles differ markedly from the lowland Amazonian territories and secondary urban centers that dominate the training distribution (Figure 2). Because subnational GDP data are not available for Bolivia, this issue does not manifest empirically for La Paz in the validation exercises. However, the Brazilian case illustrates that such boundary jurisdictions should be interpreted with caution. Their behavior reflects structural distinctiveness rather than broader model instability.

**Outlier 2: Resource-extraction transitions and structural breaks**. Pará ($R^2 = –0.078$), Rondônia ($R^2 = –0.117$), Ancash ($R^2 = –6.456$), Ayacucho ($R^2 = –1.097$), Loreto ($R^2 = –2.748$), and Pasco ($R^2 = –3.212$) exhibit negative $R^2$ under the subnational analysis, indicating that predictions perform worse than a naïve mean benchmark. Cauca ($R^2 = –18.704$) is the most extreme case and warrants separate discussion. Three patterns may explain the broader group of negative-$R^2$ units.

1. **Extractive enclaves**. Jurisdictions such as Pará, Rondônia, Loreto, and Pasco are dominated by mining and hydrocarbon activities that generate large GDP fluctuations with limited artificial illumination per unit of value added. This limitation is structural as NTL perform poorly in primary-sector, low-density contexts (Gibson et al., 2021), and the NTL–GDP elasticity declines as the primary-sector share increases (Bluhm & McCord, 2022). Even higher-quality VIIRS data struggle to capture rural extraction-based activity (Pagaduan, 2022). In these settings, output shocks driven by commodity cycles are only weakly observable through luminosity-based proxies.
2. **High-altitude Andean mining regions.** Departments such as Ancash, Ayacucho, and Pasco differ structurally from the lowland Amazonian units that dominate the training sample. Peru's mining sector is concentrated in these Andean departments (Ministerio de Energía y Minas, 2023). The spatial organization of high-altitude extraction, characterized by concentrated mine sites and dispersed spillovers, does not align with the NTL-to-GDP mapping learned from lowland agro-extractive economies.
3. **Post-conflict structural volatility.** Cauca reflects a distinct mechanism, rapid, non-monotonic structural transformation combined with illicit-economy dynamics weakly captured by satellite signals. Post-2016 economic surges followed by renewed instability (Bernal et al., 2023), expansion of coca cultivation (Nilsson & Marin, 2021), and land-use patterns difficult to distinguish via remote sensing (Murillo-Sandoval et al., 2023) produced GDP changes decoupled from NTL emissions.

By contrast, the majority of NUTS1 units exhibit $R^2$ values between 0.4 and 0.97, indicating satisfactory spatial generalization in lowland Amazonian regions characterized by stable economic structures. Colombian departments Guainía ($R^2 = 0.946$) and Guaviare ($R^2 = 0.963$) achieve the highest subnational $R^2$ in the table, reflecting the close alignment between their predominantly rural, NTL-traceable activity mix and the model's training environment.

### *5.4. Spatial Autocorrelation of Residuals*

Tables 8–10 report Moran's I Z-scores for residuals, predicted GDP, and official GDP across model specifications, estimation samples, and spatial weight matrices. Four findings emerge.

First, residual spatial autocorrelation declines monotonically from Model 1 to Model 5. Model 1, based primarily on NTL and population density, produces residual Moran Z-scores exceeding the 5 percent significance threshold in the full sample ($Z > 2.0$ from 2012 onward, peaking above 11 in 2016–17) and, to a lesser extent, in the Brazil-only sample. Model 5 exhibits improved behavior, in the Brazil-only sample, Z-scores remain within or near conventional significance bands in most years. This pattern confirms that the spatial spillover variables introduced in Models 4 and 5 effectively internalize spatial dependence that would otherwise inflate reported performance metrics. It also provides indirect validation of the LOCO

finding: the model's superior fit in structurally homogeneous contexts (Brazil only) is reflected not only in lower prediction error but also in better-behaved residual spatial structure.

Second, a pronounced anomaly appears in 2008 for Models 4 and 5. Model 4 displays residual Moran Z-scores above 38 in the full sample and above 22 in the Brazil-only sample; Model 5 shows a smaller but still substantial spike. This discontinuity could be attributable to the DMSP–VIIRS sensor transition embedded in the Chen et al. (2021a) harmonized series. The calibration procedure aligning the two sensor families introduces a temporary distortion in spatial variance around the 2007–2009 transition. Excluding 2008 leaves all other conclusions unchanged.

Third, Brazil-only residuals are systematically better behaved than pooled multi-country residuals. For Model 5, the full sample exhibits statistically significant residual autocorrelation in selected years after 2015, whereas the Brazil-only sample remains below conventional thresholds in most years. This contrast reflects structural heterogeneity across countries, which a single pooled Random Forest cannot fully internalize. The persistence of residual clustering in the multi-country sample after 2015 coincides with the Brazilian recession of 2015–16, a period of sharp structural adjustment that affected the NTL-to-GDP relationship differently across countries and introduced cross-country divergence that the spatial spillover variables cannot absorb.

Fourth, official Brazilian GDP itself displays limited subnational spatial autocorrelation within the Amazon. Moran's I Z-scores remain below conventional significance thresholds in most years. This diagnostic suggests that strong residual clustering in early specifications reflects model misspecification rather than intrinsic spatial dependence in the outcome variable. Model 5 reduces residual spatial clustering to modest levels concentrated during the 2015–16 recession, when regional structural breaks were most pronounced.

All findings are robust to alternative spatial weight matrices (W, S, B, and KNN), indicating that the results are not artefacts of spatial-definition choices.

### *5.5. Error Metrics and Cross-Country Generalization*

The choice between RMSE and MAE as performance metrics is debated in the literature (Hodson, 2022). While both should ideally be minimized, each captures various aspects of predictive performance. RMSE penalizes large deviations more heavily and is therefore more sensitive to extreme errors that arise from measurement inconsistencies, cross-country heterogeneity, or structural breaks. MAE reflects the typical magnitude of prediction errors and is more robust to outliers. In the multi-country setting examined here, higher RMSE values reflect these heterogeneous sources of noise, while MAE provides a clearer picture of average model accuracy. The differences between RMSE and MAE across specifications in Table 5 reflect precisely these dynamics.

The results also highlight a conceptual distinction between forecasting over time and generalizing across space. A model trained with multiple countries can exhibit lower temporal prediction error simply because it faces fewer out-of-sample forecasting steps under the Rolling-Origin procedure. This does not imply that such a model will predict a single country's future GDP more accurately when new years of data become available. Instead, its advantage lies in external validity. Multi-country models learn broader structural relationships that transfer more effectively to regions without data. This finding is consistent with Fen and Undavia (2022), who show that multi-country models outperform single-country models, even those with richer covariate sets, when predicting outcomes in new or data-scarce environments. This property is central

to this study because the models must generate subnational GDP estimates for Venezuela, Guyana, and Suriname, where no official data exist.

### *5.6. Spatial Redistribution and Predictor Importance*

Comparison of the consolidated model with the NTL-based GDP estimates from Chen et al. (2021a) reveals consistent spatial reallocation patterns relative to the baseline NTL-based GDP estimates. As shown in Figure 4, the model tends to redistribute GDP away from sparsely populated rural areas and toward more connected, infrastructure-intensive urban centers. This pattern is strongest in Brazil, where several NUTS3 jurisdictions near Manaus receive upward revisions, while remote municipalities with extensive forest cover receive downward adjustments. Peru displays a similar, though more spatially concentrated, pattern. Jurisdictions closer to the western urban corridor experience minor changes, whereas more remote eastern jurisdictions show larger downward revisions.

The results in Figure 5 help explain these patterns. Neighbor-based NTL variables are the most influential predictors and appear with similar importance across specifications. This finding is consistent with the spatial structure of Amazonian economic activity, where productive clusters are surrounded by large expanses of low-luminosity territory: incorporating information from neighboring cells allows the model to exploit spatial gradients that single-cell NTL measures miss entirely. This result differs from earlier NTL-based GDP approaches that relied exclusively on NTL or population as standalone predictors. The strong importance of the interaction between population and electricity consumption is consistent with recent studies that emphasize how population structure and infrastructure mediate the relationship between NTL and economic activity. Notably, the most influential predictors are proxies for economic activity, NTL, electricity consumption, and population, while geographic and environmental features play a secondary role. Feature importance reflects predictive contribution rather than structural causality and should be interpreted accordingly. The distribution of importance values nonetheless suggests that spatial context shapes GDP estimates primarily through its effects on these observable proxies, rather than through direct geographic determination.

Neighbor-based spatial gradient features, such as the mean contrast in NTL across adjacent cells, help mitigate noise and reduce outlier effects in low-light environments. This issue has been widely documented for both DMSP and early VIIRS sensors (Price et al., 2022; Hu et al., 2022). The inclusion of lagged GDP values further improves performance by capturing the historical persistence of economic activity.

Three data-quality patterns are worth noting for users of the estimates. First, predictions for 2007 display greater volatility due to inconsistencies in the official GDP series for that year, reducing their reliability relative to subsequent periods. Second, because the Rolling-Origin procedure accumulates training observations over time, the effective training sample is largest for the most recent years; this structural feature of the validation design explains the improvement in predictive accuracy visible from 2015 onward in Figure 6. Third, Brazil and Peru exhibit the most stable estimates across the panel, reinforcing the importance of data consistency for model calibration in multi-country applications.

## 6. Application and limitations of the Model and Data

### *6.1. Applications*

Reliable and regularly updated territorial data remain scarce across the Amazonia region, constraining governments' ability to design, target, and evaluate public policies. In many parts of the region, even those undergoing rapid environmental and demographic change, basic economic indicators are produced infrequently or only at NUTS1 levels, offering limited insight into local dynamics. As a result, planning decisions often rely on anecdotal evidence or qualitative assessments rather than systematic quantitative analysis (González Saldarriaga et al., 2025).

The gridded GDP dataset produced in this study does not replace official statistics, but it provides a valuable intermediate tool for territorial analysis. By integrating nighttime activity, spatial spillovers, and a temporally consistent estimation approach, the dataset enables a clearer understanding of local economic patterns in remote or data-scarce environments. Its applications span multiple policy domains.

First, data enables detailed measurement of spatial inequality. Traditional subnational GDP masks substantial intraregional variation. The 5 × 5 km grid supports the computation of spatial Gini coefficients, localized inequality indices, and growth incidence measures, allowing policymakers to identify pockets of deprivation or emerging economic hubs that would be invisible at aggregated scales. The consistent temporal coverage also facilitates the study of how economic activity evolves alongside demographic and environmental changes.

Second, the dataset strengthens infrastructure planning and evaluation. The spatial distribution of GDP can be linked to road improvements, electrification programs, navigable river upgrades, and digital connectivity projects to assess their economic footprint. Because the model incorporates neighboring effects, it can identify whether the effects of infrastructure investments extend into adjacent areas, a critical question in a region where connectivity is uneven and transport networks are sparse.

Third, the data supports environmental and land-use policy analysis. Many conservation, restoration, and land-management decisions require an understanding of how economic activity interacts with deforestation fronts, fire incidence, protected areas, and biodiversity corridors. The GDP grid allows integrated spatial analysis of economic pressures and environmental change, helping quantify opportunity costs, identify high-risk zones, and support the design of compensation or transition mechanisms for affected communities.

Fourth, the dataset enhances disaster risk management and climate-adaptation planning. By providing granular annual estimates of local economic activity, the grid makes it possible to quantify exposure to floods, droughts, fires, and extreme climate events, and to track recovery trajectories across affected areas. This can inform the prioritization of adaptation investments and the targeting of post-disaster support.

Fifth, the data provides new insights into urbanization and settlement dynamics. The rapid growth of cities such as Manaus, Santarém, and Portel in Brazil has reshaped economic geography in the region. The GDP grids make it possible to study how activity concentrates in urban cores, how secondary cities emerge, and how economic activity diffuses along river and road systems, patterns that are not captured by national statistical reporting.

Finally, the framework contributes to statistical system strengthening. Countries with incomplete or irregular subnational GDP series may use model-based estimates to interpolate missing years, benchmark

official figures, and support gradual improvements in territorial statistical capacity. Beyond Amazonia, the methodology is transferable to other data-limited and environmentally sensitive regions, including the Congo Basin and interior Southeast Asia.

### *6.2. Future Research*

Future research can build on this framework by refining the measurement of economic activity in rural and forested environments and by expanding the set of spatial predictors available for modeling development dynamics. A priority is the integration of higher-resolution remote-sensing variables, such as vegetation indices, land-cover transitions, and indicators of ecosystem service intensity, which can capture forms of economic activity that NTL alone do not detect. This is particularly relevant for agricultural production, forest-based livelihoods, and natural-resource extraction, as well as for estimating the broader economic value of the region's ecological assets in the spirit of Hanusch (2023). In addition, emerging datasets such as the IDB's agricultural productivity data for Latin America (Salazar et al., 2025) offer promising opportunities to incorporate sector-specific signals that can further improve the characterization of rural economic activity.

A second avenue is the development of more sophisticated measures of spatial accessibility. In the Amazonia region, connectivity is a fundamental constraint on economic opportunities: vast distances, sparse road networks, and the predominance of river transport shape patterns of market access, mobility, and service provision. Future work could integrate multimodal transport networks, including road, river, and digital connectivity, into accessibility indices that quantify how easily households and firms can reach markets and services. Existing datasets, such as the GIS-based transport layers of Rocha et al. (2023) for the Amazonia region, could be combined with methodologies like those in Duran-Fernandez and Santos (2014) to generate these indices. Incorporating such measures into the modeling framework may improve predictions in isolated regions where economic potential is shaped more by accessibility than by local production alone.

Advances in deep learning also offer promising opportunities for model development. Convolutional neural networks and spatiotemporal architectures may further enhance predictive accuracy by leveraging the fine spatial structure of VIIRS imagery, land-cover mosaics, and other geospatial layers. Hybrid approaches that embed economically motivated structural constraints within machine-learning models could improve interpretability and allow the estimation of more nuanced non-linear relationships, especially in heterogeneous terrain. As more countries release disaggregated GDP and demographic data, country-specific model variants should also be explored to assess how generalizable the multi-country framework is.

The framework lends itself to forward-looking scenario analysis. By re-estimating the model under alternative assumptions, such as electrification expansion, infrastructure upgrades, river-transport improvements, migration flows, or land-use transitions, researchers could generate counterfactual trajectories of local economic activity under alternative assumptions. While these are proxy-based rather than structural forecasts, they can nonetheless provide relevant insights for long-term planning in regions experiencing rapid environmental, climatic, and demographic change.

Finally, from a methodological perspective, the counterintuitive performance gap between single-country and multi-country models under temporal cross-validation opens a promising methodological research

agenda. It highlights the need for evaluation metrics and validation schemes that jointly capture temporal forecasting accuracy and spatial generalization capacity, an area that remains unexplored in satellite-based GDP literature.

### *6.3. Limitations*

The framework has limitations that should guide interpretation of the results and inform future research.

The first concerns the uneven availability of subnational GDP data across countries. Within the study area, only Brazil, Peru, and Colombia provide subnational GDP statistics, and these differ in temporal depth and geographic granularity. Brazil supplies annual municipal-level data for 2002–2019; Peru provides district-level data for 1993–2018; and Colombia offers municipio-level data only for 2011–2019. The short Colombian panel, combined with the absence of subnational GDP data for Bolivia, Ecuador, Venezuela, Guyana, Suriname, and French Guiana, implies that the model is calibrated primarily on the southern and northern portions of the Basin and extrapolated to the remaining territory. Consequently, the estimated relationships may disproportionately reflect the structural and temporal characteristics of countries with longer and richer data coverage.

A second limitation relates to the two-stage estimation strategy. The Random Forest is trained at the NUTS3 level, and predicted values are subsequently downscaled to the 5 × 5 km grid using a mass-preserving allocation rule. We explored direct grid-level estimation to avoid this top-down procedure, but training over millions of grid cells across multiple years proved computationally infeasible on available hardware. The adopted approach follows the logic of Chen et al. (2021a), who implement a similar framework at 1 km resolution in a global setting. Although this ensures internal consistency between administrative and grid-level estimates, direct end-to-end grid-based estimation would eliminate the intermediate aggregation layer and may improve spatial precision. As computational capacity and training architecture expand (e.g., Rossi-Hansberg and Zhang, 2025), fully grid-level estimation represents a natural extension of the framework.

A third limitation concerns population data quality. Recent evidence (Láng-Ritter et al., 2025) documents substantial underestimation of rural populations in widely used global gridded datasets. Because GHS-POP is a central input to the model, systematic undercounting of rural settlements may translate into downward biases in predicted GDP for remote or sparsely populated areas. Future research should prioritize the integration of improved demographic datasets, country-specific census layers, or alternative high-frequency proxies derived from mobile-phone activity, settlement footprints, or very high-resolution imagery.

The robustness exercises reported in Section 5 clarify an additional structural constraint: spatial generalization is conditional on the characteristics of the training countries. The $R^2$ values of 0.957–0.958 reported in Table 5 reflect temporal cross-validation performance within the estimation sample and should be interpreted accordingly. They do not imply equivalent accuracy for territories structurally distinct from Brazil, Peru, and Colombia. The LOCO and subnational robustness exercises demonstrate that predictive performance deteriorates when the model is required to extrapolate to an unseen national context, particularly when the excluded country is Brazil, which dominates the training sample in both observation count and economic scale. The practical use of the gridded estimates should therefore be conditioned on structural comparability with the countries informing model calibration.

The subnational cross-validation further identifies a subset of NUTS1 units where performance is systematically weak. These include regions characterized by large-scale extractive enclaves (e.g., Pará,

Pasco, Loreto), rapid structural transitions (e.g., Cauca), and atypical administrative economies (e.g., Distrito Federal). In these cases, prediction errors reflect structural mismatch rather than random noise. Users should exercise caution in interpreting estimates for such regions and, where possible, supplement model outputs with sector-specific administrative data.

Finally, all official GDP data were originally reported in nominal local-currency units and converted into real 2017 USD for cross-country comparability. This required downscaling national GDP deflators to the subnational level using proportional adjustments derived from Chen et al. (2021a). While this procedure produces a harmonized regional dataset, it rests on assumptions that may not fully capture subnational price variation or inflation heterogeneity. Access to finer-grained local deflators or subnational price indices would materially improve cross-country comparability and temporal consistency.

## 7. Final remarks

This study proposes a new methodological framework for estimating subnational GDP in data-scarce environments by integrating machine learning, NTL, spatial spillovers, and temporal cross-validation. Applied to the Amazonia region, one of the most complex and least measured territories in the world, the framework delivers consistent predictive gains over traditional luminosity-based approaches and provides a coherent strategy for generating harmonized, temporally aligned GDP estimates across seven countries. By combining NTL intensity, electricity consumption, population, lagged structural variables, and spatial-interaction features, the model captures a richer and more stable signal of economic activity than any single proxy can provide. The robustness exercises define the boundaries of this reliability through a structured validation hierarchy that progresses from temporal forecasting within the training sample to spatial generalization across held-out countries, and finally to subnational extrapolation and residual diagnostic testing.

A central methodological contribution of the paper lies in distinguishing two predictive tasks that are often conflated in the satellite-based GDP literature: temporal forecasting accuracy and spatial generalization. Although literature treats these dimensions as interchangeable, the results demonstrate that they are fundamentally distinct and require separate evaluation environments. Standard error metrics, when applied to panels with heterogeneous time-series length, can generate misleading inferences regarding comparative model performance. The divergence between temporal and spatial cross-validation illustrates this directly. A model that achieves an $R^2$ of 0.957 under rolling-origin temporal evaluation yields $R^2$ values between 0.15 and 0.41 under LOCO testing, a difference that would remain obscured if only a single-sample performance statistic were reported.

The evidence regarding temporal stability is clear. Rolling-origin temporal cross-validation produces $R^2$ values above 0.94 for GDP levels in every year from 2014 onward, with a full-period average of 0.957. The model tracks year-to-year variation within the training countries with high fidelity and exhibits no indication of systematic overfitting. None of the robustness exercises modifies this conclusion.

The evidence regarding spatial generalization is more nuanced. The proposition that multi-country estimation enhances spatial generalization must be interpreted considering the LOCO results. When trained on two countries and evaluated on the third, the model generalizes with moderate accuracy to Colombia ($R^2 = 0.372$) and Peru ($R^2 = 0.412$) yet performs poorly when trained exclusively on Colombia and Peru and evaluated on Brazil ($R^2 = 0.146$). This asymmetry reflects Brazil's structural dominance in the training sample, both in observation count and economic scale, rather than a failure of the multi-country framework itself. The model performs best in contexts structurally like the Brazilian Amazonia. Confidence in extrapolated predictions should therefore vary accordingly: it is highest for Brazilian Amazonian *municipios*, moderate for Colombian and Peruvian departments and districts, and more qualified for countries entirely outside the training sample. This graded structure of confidence is not unique to this application but is inherent to any machine-learning estimator trained on observational spatial data. Making that hierarchy explicit constitutes an important contribution to the satellite-based GDP literature.

Predictions for Bolivia, Ecuador, Venezuela, Guyana, and Suriname, should be understood as extrapolations conditioned on the structural comparability of those countries' Amazonian economies with the training countries. For Amazonian lowland territories with agro-extractive economic profiles, such as

much of Bolivia and Ecuador, this assumption is plausible, and the gridded estimates can serve as inputs to regional planning, albeit with wider uncertainty bands than in the training countries. For Venezuela and the Guianas, where institutional structures, data-generating processes, and macroeconomic dynamics differ more substantially from the estimation sample, extrapolation is more speculative and should be accompanied by explicit acknowledgement of this limitation. In all cases, the appropriate use of these estimates is as a first-order approximation to be refined as subnational administrative data become available, rather than as a substitute for national accounts.

The spatial autocorrelation diagnostics further clarify model performance. Moran's analysis shows that Model 5 reduces residual spatial dependence relative to simpler specifications, validating the inclusion of spatial spillover variables. A pronounced anomaly in 2008 reflects the DMSP–VIIRS sensor transition embedded in the harmonized Chen et al. (2021a) series, which introduces a temporary discontinuity in spatial variance. Users requiring sensor-consistent estimates may restrict analysis to the post-2009 period or, more conservatively, to the VIIRS-native period beginning in 2012.

Taken together, the temporal cross-validation, LOCO, subnational cross-validation, and spatial autocorrelation analyses define a coherent validation hierarchy. The model produces highly reliable estimates within the temporal and geographic scope of its training data, and its accuracy degrades in a predictable and interpretable manner as we move further from the training distribution, whether across time or across space, in countries outside the estimation sample or structurally atypical subnational units. By characterizing this degradation explicitly rather than reporting a single aggregate performance metric, the paper offers a framework for transparent performance disclosure in satellite-based economic measurement, a standard that becomes increasingly important as such estimates enter policy and planning workflows.

## 6. Bibliography


[1.] Bernal, C., Prem, M., Vargas, J. F., & Ortiz, M. (2024). Peaceful entry: Entrepreneurship dynamics during Colombia's peace agreement. *Journal of Development Economics, 166*, 103119. https://doi.org/10.1016/j.jdeveco.2023.103119

[2.] Bickenbach, F., Bode, E., Nunnenkamp, P., & Söder, M. (2016). Night lights and regional GDP. *Review of World Economics, 152*(2), 425–447. https://doi.org/10.1007/s10290-016-0246-0

[3.] Bluhm, R., & McCord, G. C. (2022). What can we learn from nighttime lights for small geographies? Measurement errors and heterogeneous elasticities. *Remote Sensing, 14*(5), 1190. https://doi.org/10.3390/rs14051190

[4.] Bundervoet, T., Maiyo, L., & Sanghi, A. (2015). *Bright lights, big cities: Measuring national and subnational economic growth in Africa from outer space, with an application to Kenya and Rwanda* (World Bank Policy Research Working Paper No. 7461). World Bank. https://doi.org/10.1596/1813-9450-7461

[5.] Carioli, A., Schiavina, M., Freire, S., & MacManus, K. (2023). *GHS-POP R2023A – GHS population grid multitemporal (1975–2030)* [Dataset]. European Commission, Joint Research Centre. https://doi.org/10.2905/2FF68A52-5B5B-4A22-8F40-C41DA8332CFE

[6.] Chen, J., & Gao, M. (2021a). *Global 1 km × 1 km gridded revised real gross domestic product and electricity consumption during 1992–2019: Real GDP* [Dataset]. https://doi.org/10.6084/m9.figshare.17004523

[7.] Chen, J., & Gao, M. (2021b). *Global 1 km × 1 km gridded revised real gross domestic product and electricity consumption during 1992–2019: Electricity consumption* [Dataset].https://doi.org/10.6084/m9.figshare.17004523.v1

[8.] Chen, J., Gao, M., Cheng, S., Hou, W., Song, M., Liu, X., & Liu, Y. (2022). Global 1 km × 1 km gridded revised real gross domestic product and electricity consumption during 1992–2019 based on calibrated nighttime light data. *Scientific Data, 9*(1), Article 202. https://doi.org/10.1038/s41597-022-01322-5

[9.] Chen, X., & Nordhaus, W. D. (2011). Using luminosity data as a proxy for economic statistics. *Proceedings of the National Academy of Sciences, 108*(21), 8589–8594. https://doi.org/10.1073/pnas.1017031108

[10.] Chen, Z., Yu, B., Yang, C., Zhou, Y., Yao, S., Qian, X., Wang, C., Wu, B., & Wu, J. (2020). *An extended time-series (2000–2023) of global NPP-VIIRS-like nighttime light data (Version 5)* [Dataset]. Harvard Dataverse. https://doi.org/10.7910/DVN/YGIVCD

[11.] da Rocha Bragion, G., Monteiro, A. M. V., & Amaral, S. (2019). Exploring VIIRS-NPP night-time light data in the Amazon rainforest. *Remote Sensing Applications: Society and Environment, 15*, 100221. https://doi.org/10.3390/rs14133126

[12.] Delafontaine, M., Nolf, G., Van de Weghe, N., Antrop, M., & De Maeyer, P. (2009). Assessment of sliver polygons in geographical vector data. *International Journal of Geographical Information Science, 23*(6), 719–735. https://doi.org/10.1080/13658810701694838

[13.] Donaldson, D., & Storeygard, A. (2016). The view from above: Applications of satellite data in economics. *Journal of Economic Perspectives, 30*(4), 171–198. https://doi.org/10.1257/jep.30.4.171

[14.] Duran-Fernandez, R., & de Carvalho Coutinho, T. (2025). Amazonia: A puzzle between development and sustainability? *Latin American Policy, 16*(4), e800. https://doi.org/10.1111/lamp.70030

[15.] Duran-Fernandez, R., & Santos, G. (2014). A regional model of road accessibility in Mexico: Accessibility surfaces and robustness analysis. *Research in Transportation Economics, 46*, 55–69. https://doi.org/10.1016/j.retrec.2014.09.017

[16.] European Union. (2024). *Council Regulation (EC) No. 1059/2003 of 26 May 2003 on the establishment of a common classification of territorial units for statistics (NUTS) (as amended up to 1 January 2024)*. EUR-Lex. https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:02003R1059-20240101

[17.] Feenstra, R. C., Inklaar, R., & Timmer, M. P. (2015). The next generation of the Penn World Table. *American Economic Review, 105*(10), 3150–3182. https://doi.org/10.1257/aer.20130954

[18.] Fen, C., & Undavia, S. (2022). Improving macroeconomic model validity and forecasting performance with pooled country data using structural, reduced-form, and neural network models. *arXiv*. https://arxiv.org/abs/2203.06540

[19.] Fujita, M., Krugman, P., & Venables, A. (1999). *The spatial economy: Cities, regions, and international trade*. MIT Press.

[20.] Galimberti, J. K. (2020). Forecasting GDP growth from outer space. *Oxford Bulletin of Economics and Statistics, 82*(4), 697–722. https://doi.org/10.1111/obes.12361

[21.] Gibson, J., Olivia, S., Boe-Gibson, G., & Li, C. (2021). Which night lights data should we use in economics, and where? *Journal of Development Economics, 149*, 102602. https://doi.org/10.1016/j.jdeveco.2020.102602

[22.] González Saldarriaga, S., García Sánchez, V., & Durán-Fernández, R. (2025). *Regiones de frontera de la Amazonía: Hacia un desarrollo productivo sostenible*. Inter-American Development Bank. https://publications.iadb.org/es/regiones-de-frontera-de-la-amazonia-hacia-un-desarrollo-productivo-sostenible

[23.] Hanusch, M. (2023). *A balancing act for Brazil's Amazonian states: An economic memorandum*. World Bank. https://bit.ly/BalancingActFullEN

[24.] Henderson, J. V., Storeygard, A., & Weil, D. N. (2012). Measuring economic growth from outer space. *American Economic Review, 102*(2), 994–1028. https://doi.org/10.1257/aer.102.2.994

[25.] Hodson, T. O. (2022). Root mean square error (RMSE) or mean absolute error (MAE): When to use them or not. *Geoscientific Model Development, 15*, 5481–5487. https://doi.org/10.5194/gmd-15-5481-2022

[26.] Hu, Y., & Yao, J. (2022). Illuminating economic growth. *Journal of Econometrics, 228*(2), 359–378. https://doi.org/10.1016/j.jeconom.2021.10.007

[27.] IJSN (2024). *PIB municipal – Cadernos*. Instituto Jones dos Santos Neves https://ijsn.es.gov.br/publicacoes/cadernos/pib-municipal

[28.] Jean, N., Burke, M., Xie, M., Davis, W. M. A., Lobell, D. B., & Ermon, S. (2016). Combining satellite imagery and machine learning to predict poverty. *Science, 353*(6301), 790–794. https://doi.org/10.1126/science.aaf7894

[29.] Jin, Y., Ge, Y., Fan, H., Li, Z., Liu, Y., & Jia, Y. (2023). Mapping gross domestic product distribution at 1 km resolution across Thailand using the random forest area-to-area regression kriging model. *ISPRS International Journal of Geo-Information, 12*(12), 481. https://doi.org/10.3390/ijgi12120481

[30.] Khachiyan, A., Thomas, A., Zhou, H., Hanson, G. H., Cloninger, A., Rosing, T., & Khandelwal, A. K. (2022). Using neural networks to predict microspatial economic growth. American Economic Review: Insights, 4(4), 491–506. https://10.1257/aeri.20210422

[31.] Láng-Ritter, J., Keskinen, M., & Tenkanen, H. (2025). Global gridded population datasets systematically underrepresent rural population. *Nature Communications, 16*(1), 2170. https://doi.org/10.1038/s41467-025-56906-7

[32.] Martinez, L. R. (2022). How much should we trust the dictator's GDP growth estimates? *Journal of Political Economy, 130*(10), 2731–2769. https://doi.org/10.1086/720458

[33.] Mellander, C., Lobo, J., Stolarick, K., & Matheson, Z. (2015). Night-time light data: A good proxy measure for economic activity? *PLOS ONE, 10*(10), e0139779. https://doi.org/10.1371/journal.pone.0139779

[34.] Ministerio de Energía y Minas del Perú. (2023). *Anuario minero 2022*. Gobierno del Perú. https://www.gob.pe/institucion/minem/informes-publicaciones/4326371-anuario-minero-2022

[35.] Murillo-Sandoval, P. J., Kilbride, J., Tellman, E., Wrathall, D., Van Den Hoek, J., & Kennedy, R. E. (2023). The post-conflict expansion of coca farming and illicit cattle ranching in Colombia. *Scientific Reports, 13*(1), 1965. https://doi.org/10.1038/s41598-023-28918-0

[36.] Nilsson, M., & Marín, L. G. (2021). Colombia's program to substitute crops used for illegal purposes: Its impact on security and development. *Journal of Intervention and Statebuilding, 15*(3), 309–326. https://doi.org/10.1080/17502977.2021.1921546

[37.] Pagaduan, J. A. (2022). Do higher-quality nighttime lights and net primary productivity predict subnational GDP in developing countries? Evidence from the Philippines. *Asian Economic Journal, 36*(3), 288–317. https://doi.org/10.1111/asej.12278

[38.] Pérez-Sindín, X. S., Chen, T.-H. K., & Prishchepov, A. V. (2021). Are night-time lights a good proxy of economic activity in rural areas in middle- and low-income countries? *Remote Sensing Applications: Society and Environment, 24*, 100647. https://doi.org/10.1016/j.rsase.2021.100647

[39.] Potapov, P., Turubanova, S., Hansen, M. C., Tyukavina, A., Zalles, V., Khan, A., Song, X.-P., Pickens, A., Shen, Q., & Cortez, J. (2021). Global maps of cropland extent and change show accelerated cropland expansion in the twenty-first century. *Nature Food, 2*, 720–730. https://doi.org/10.1038/s43016-021-00429-z

[40.] Price, N., & Atkinson, P. M. (2022). Global GDP prediction with night-lights and transfer learning. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 15*, 7128–7138. https://doi.org/10.1109/JSTARS.2022.3200754

[41.] RAISG. (2020). *Amazon in numbers*. https://www.raisg.org/en/infographic/

[42.] RAISG. (2023a). *Illegal mining* [Dataset]. https://www.raisg.org/en/download/illegal-mining/

[43.] RAISG. (2023b). *Natural protected areas* [Dataset]. https://www.raisg.org/en/download/natural-protected-areas/

[44.] RAISG. (2023c). *Roads* [Dataset]. https://www.raisg.org/en/download/roads-2/

[45.] RAISG. (2024a). *Deforestation total RAISG* [Dataset]. https://www.raisg.org/en/download/deforestation-total-raisg/

[46.] RAISG. (2024b). *Hydroelectric plants* [Dataset]. https://www.raisg.org/en/download/zonas-mineras/

[47.] RAISG. (2024c). *Mining blocks* [Dataset]. https://www.raisg.org/en/download/zonas-mineras/

[48.] RAISG. (2024d). *Oil* [Dataset]. https://www.raisg.org/en/download/zonas-mineras/

[49.] Roberts, D. R., Bahn, V., Ciuti, S., Boyce, M. S., Elith, J., Guillera-Arroita, G., & Dormann, C. F. (2017). Cross-validation strategies for data with temporal, spatial, hierarchical, or phylogenetic structure. *Ecography, 40*(8), 913–929. https://doi.org/10.1111/ecog.02881

[50.] Rocha, T. A. H., Silva, L. L., Wen, F. H., et al. (2023). River dataset as a potential fluvial transportation network for healthcare access in the Amazon region. *Scientific Data, 10*, Article 188. https://doi.org/10.1038/s41597-023-02085-3

[51.] Rossi-Hansberg, E., & Zhang, J. (2025). *Local GDP estimates around the world* (NBER Working Paper No. 33458). National Bureau of Economic Research. https://doi.org/10.3386/w33458

[52.] Salazar, L., Schling, M., De Salvo, C. P., & Martel, P. (2025). *Productividad agrícola en América Latina y el Caribe: Qué sabemos y hacia dónde vamos*. Inter-American Development Bank. https://doi.org/10.18235/0013811

[53.] Sekulic, A., Kilibarda, M., Heuvelink, G. B. M., Nikolic, M., & Bajat, B. (2020). Random forest spatial interpolation. *Remote Sensing, 12*(10), 1687. https://doi.org/10.3390/rs12101687

[54.] World Wildlife Fund. (2006). *South America: Void-filled digital elevation model, 15s resolution* [Dataset]. U.S. Geological Survey.

[55.] World Wildlife Fund. (2023). *Inside the Amazon*. https://wwf.panda.org/discover/knowledge_hub/where_we_work/amazon/about_the_amazon/

[56.] Yeh, C., Perez, A., Driscoll, A., Azzari, G., Tang, Z., Lobell, D., & Burke, M. (2020). Using publicly available satellite imagery and deep learning to understand economic well-being in Africa. *Nature Communications, 11*, 2583. https://doi.org/10.1038/s41467-020-16185-w

[57.] Zhao, M., Cheng, C., Zhou, Y., Li, X., Shen, S., & Song, C. (2022). A global dataset of annual urban extents (1992–2020) from harmonized nighttime lights. *Earth System Science Data, 14*, 517–534. https://doi.org/10.5194/essd-14-517-2022

**Annex 1. LCU Standardization**

To maintain consistency across currencies and sources, all official GDP figures in the study were standardized to real 2017 USD.

First, the fraction of each country's Amazonian region's GDP relative to national GDP was calculated using the calibrated NTL-based GDP data from Chen et al. (2021a). This fraction was then multiplied by the Penn World Table's national GDP estimates for each country to obtain an estimate of Amazonian GDP in real 2017 USD (Feenstra et al., 2015). This process is formalized in Equation A1.

**Equation A1.** Amazonian GDP Estimation

$$GDP\_NUTS0\_AMZ_{i,t} = \sum_{k \in A} \frac{GDP\_NUTS3\_Chen_{i,j,k,t}}{GDP\_NUTS0\_Chen_{i,t}} GDP\_NUTS0\_Penn_{i,t}$$

Where:

*i* represents the national level (NUTS0)

*j* represents the second level subnational jurisdiction (NUTS1)

*k* represents the local level subnational jurisdiction (NUTS3)

*t* represents the year

*A* is the Amazonian region

Thus, by summing the GDP proportions of municipalities *k* located in the Amazon and multiplying this fraction by the Penn World Table's national GDP, we obtain the total value of the Amazonian economy for each country *i* and year *t*.

Subsequently, the official NUTS2 GDP data, originally reported in LCU, was converted into standardized real USD by multiplying the previously calculated Amazonian GDP in real USD by the proportion of the NUTS2 unit's LCU GDP over the country's Amazonian GDP in LCU. This is formalized in Equation A2:

**Equation A2.** NUTS2 Standardized GDP Conversion

$$GDP\_NUTS3\_USD2017_{i,j,k,t} = GDP\_NUTS0\_AMZ_{i,t} \frac{GDP\_NUTS3\_LCU_{i,j,k,t}}{\sum_{k \in A} GDP\_NUTS3\_LCU_{i,j,k,t}}$$

This two-step conversion ensures full comparability between official GDP figures and the machine learning predictions, eliminating inconsistencies caused by currency fluctuations or mixed monetary units.

**Table 1 Subnational Divisions in Amazonia**

| NUTS0 | NUTS1 | NUTS2 | NUTS3 |
|---|---|---|---|
| Bolivia | *Departamento* | *Provincia** | *Municipios* |
| Brazil | *Estado* | | *Municipios** |
| Colombia | *Departamento* | | *Municipios** |
| Ecuador | *Provincia* | *Canton** | *Parroquias* |
| Guyana | *Regions* | | *Neighborhood Councils** |
| Peru | *Departamento* | *Provincia* | *Distritos** |
| Suriname | *Distrikten* | | *Ressort ** |
| Venezuela | *State* | | *Municipios** |

* Asterisks indicate the administrative level at which GDP forecasts are produced. Model training uses subnational GDP data only for countries where such data are available.

**Table 2 GIS Data**

| Data | Type of Data | Time Period | Resolution | Source |
|---|---|---|---|---|
| GPD | Raster of real GDP estimated with night lights in real USD | 2000-2019 | 1 km x 1 km | (Chen et al., 2020) |
| Population | Raster of population in number of inhabitants | 2005, 2010 | 100m x 100m | Carioli et al. (2023 |
| State and departmental roads | Location represented by a line vector | 2023 | - N.A. | (RAISG, 2023c) |
| Roads | Location represented by a line vector | 2021 | - | (Rocha et al., 2023) |
| Nighttime Luminosity | Raster of radiance of emitted light (nanowatts per square centimeter per steradian) | 2000 - 2023 | 500m x 500m | (Chen et al., 2020) |
| Electricity Consumption | Raster of radiance of emitted light (watts per square centimeter per steradian) | 1990 - 2019 | 1km x 1km | (Chen et al., 2021b) |
| Urban Area | Raster representing percentage of urban space per pixel | 1992 - 2020 | 1km x 1km | (Zhao et al., 2022) |
| Harvest Area | Raster representing percentage of harvested space per pixel | 2003, 2007, 2011, 2015, 2019 | 30m x 30m | (Potapov et al., 2021) |
| Elevation | Raster with elevation in meters | 2006 | 15 arcseconds | (WWF, 2006) |
| Illegal Mining | Location represented by a polygon vector | 2023 | N.A. | (RAISG, 2023a) |
| Protected Natural Areas | Location represented by a polygon vector | 2023 | N.A. | (RAISG, 2023b) |
| Hydroelectric Plants | Location represented by a polygon vector | 2024 | N.A. | (RAISG, 2024b) |
| Legal Mining Areas | Location represented by a polygon vector | 2024 | N.A. | (RAISG, 2024c) |
| Oil Extraction Zones | Location represented by a polygon vector | 2024 | N.A. | (RAISG, 2024d) |

**Table 3 Summary of Constructed Data**

| Variable | Formula | Description |
|---|---|---|
| GDP Density | $\frac{GDP_{i,j,k,t}}{Area\ (km^2)_{i,j,k,t}}$ | GDP divided by $km^2$ |
| Population Density | $\frac{Population_{i,j,k,t}}{Area\ (km^2)_{i,j,k,t}}$ | Population divided by $km^2$ |
| GDP 3 year moving average | N.A. | The average of the GDP from current and last 2 years. |
| Population Density 3 year moving average | N.A. | The average of the population density from current and last 2 years. |
| Electric Consumption 3 year moving average | N.A. | The average of the electric consumption from current and last 2 years. |
| GDP NUTS2/NUTS1 proportion | $\frac{GDP_{i,j,k,t}}{GDP_{i,j,t}}$ | Proportion of the GDP that a certain NUTS2 entity represents of the whole NUTS1 |
| GDP percent growth | $\frac{GDP_{i,j,k,t} - GDP_{i,j,k,t-1}}{GDP_{i,j,kt-1}}$ | Percentage change of the GDP from last to current year |
| Population density growth | $\frac{Pop\ Density_{i,j,k,t} - Pop\ Density_{i,j,k,t-1}}{Pop\ Density_{i,j,k,t-1}}$ | Percentage change of the population density from last to current year |
| Deforestation growth | $\frac{Deforest_{i,j,k,t} - Deforest_{i,j,k,t-1}}{Deforest_{i,j,k,t-1}}$ | Percentage change of deforestation from last to current year |
| Night time light growth | $\frac{NTL_{i,j,k,t} - NTL_{i,j,k,t-1}}{NTL_{i,j,k,t-1}}$ | Percentage change of night time light intensity from last to current year |
| Electric consumption growth | $\frac{EC_{i,j,k,t} - EC_{i,j,k,t-1}}{EC_{i,j,k,t-1}}$ | Percentage change of electric consumption from last to current year |
| Elevation differential | $Max\ Elev_{i,j,k} - Min\ Elev_{i.j.k}$ | Gets the elevation difference in the country |
| ANP per Area | $\frac{\sum_{c=1}^{N} ANP_{i,j,k,c}}{Area\ (km^2)_{i,j,k}}$ | Calculates the amount of Natural Protected Areas by square kilometer space. |
| Petrol per Area | $\frac{\sum_{c=1}^{N} Petrol\ station_{i,j,k,c}}{Area\ (km^2)_{i,j,k}}$ | |
| Roads per Area | $\frac{Road\ length_{i,j,k}}{Area\ (km^2)_{i,j,k}}$ | |
| Hydroelectric Plant per Area | $\frac{\sum_{c=1}^{N} Hydro\ Plant_{i,j,k,c}}{Area\ (km^2)_{i,j,k}}$ | |
| Mountainous | N.A. | Dummy variable if the area surpasses the defined altitude for mountainous areas, being 2012 meters over sea level. |
| Crop / Deforestation | Crop * Deforestation | Interaction term between percent of crop land and deforestation |
| Population / Electric Consumption | Population * EC | Interaction term between population and electricity consumption |
| Population / Night Time Lights | Population * NTL | Interaction term between population and night time lights |

| Variable | Formula | Description |
|---|---|---|
| Population / Urban | Population * Urban | Interaction term between population and percentage of urban area |
| Night time light / Urban | NTL * Urban | Interaction term between night time lights and percentage of urban area |
| Night time lights / electric consumption | NTL * EC | Interaction term between night time lights and electricity consumption |
| Deforestation lag / population | N.A. | Interaction term between Lag of deforestation and population |
| Deforestation lag / urban | N.A. | Interaction term between lag of deforestation and percentage of urban area |
| Road count / Urban | N.A. | Interaction term between Amount of roads and percentage of urban area |
| Round length / Urban | N.A. | Interaction term between Total length of roads and percentage of urban area |
| Nighttime lights squared | $NTL^2$ | Squared value of night time lights |
| Electricity consumption squared | $EC^2$ | Squared value of electric consumption |

Neighbor Variables

| Variable | Variables Included | Description |
|---|---|---|
| Max Neighbor | - NTL<br>- EC<br>- Pop Density<br>- Deforestation<br>- GDP NUTS2<br>- GDP Density | Maximum value of the variable of the units neighbors |
| Minimum Neighbor | | Minimum value of the variable of the units neighbors |
| Mean Neighbor | | Mean value of the variable of the units neighbors |
| Max Difference Neighbor | | Difference between the unit and the maximum value of its neighbors |
| Minimum Difference Neighbor | | Difference between the unit and the minimum value of its neighbors |
| Mean Difference Neighbor | | Difference between the unit and the mean value of its neighbors |
| Relative Difference Maximum | | Percentage difference between the unit and the maximum value of its neighbors |
| Relative Difference Minimum | | Percentage difference between the unit and the minimum value of its neighbors |
| Relative Difference Mean | | Percentage difference between the unit and the mean value of its neighbors |

**Note:** *i* represents NUTS0, *j* represents NUTS1, *k* represents NUTS2, l represents cell, *t* represents the year

**Table 4 Model Description**

| Group | Description | Variables | Model 1 (1) | Model 2 (2) | Model 3 (3) | Model 4 (4) | Model 5 (5) | Model 6 (6) |
|---|---|---|---|---|---|---|---|---|
| *PAC* | Panel Alignment Controls | Chen GDP | x | x | | | x | x |
| | | Proportion of NUTS2/NUTS1 population | x | x | | | x | x |
| | | Brazil Dummy | x | x | | | x | x |
| | | Colombia dummy | x | x | | | x | x |
| | | Peru dummy | x | x | | | x | x |
| *STR* | Structural Baseline Variables | Population | x | x | x | x | x | x |
| | | Population density | x | x | x | x | x | x |
| | | Deforestation | x | x | x | x | x | x |
| $STR^D$ | Structural Dynamics | Population \| Moving average 2 lags | | | x | x | x | x |
| | | Population density \| 1st lag | | | x | x | x | x |
| | | Population density \| 2nd lag | | | x | x | x | x |
| | | Population density \| moving average 2 lag | | | x | x | x | x |
| | | Deforestation \| 1st lag | | | x | x | x | x |
| | | Deforestation \| growth | | | x | x | x | x |
| $STR^S$ | Structural Neighboring | Population \| Neighboring | | | | /1 | /1 | /1 |
| | | Population Density \| Neighboring | | | | x | x | x |
| | | Deforestation \| Neighboring | | | | x | x | x |
| *NTA* | Nighttime Activity Indicators | Night lights | | | x | x | | x |
| | | Electric consumption | | | x | x | | x |
| $NTA^S$ | Nighttime Activity Spillovers | NTL \| Neighboring | | | | x | | x |
| | | Electric consumption \| Neighboring | | | | x | | x |
| *ECO* | Economic Level & Trend Indicators | Chen GDP \| 1st lag | | | | | x | x |
| | | Chen GDP \| 2nd lag | | | | | x | x |
| | | Chen GDP \| moving average 3 lags | | | | | x | x |
| | | Chen GDP \| per sq km | | | | | x | x |
| | | Chen GDP \|NUTS2 growth | | | | | x | x |
| | | Chen GDP \|NUTS2 percentage increase | | | | | x | x |

| Group | Description | Variables | Model 1 (1) | Model 2 (2) | Model 3 (3) | Model 4 (4) | Model 5 (5) | Model 6 (6) |
|---|---|---|---|---|---|---|---|---|
| | | Chen GDP | | | | | x | x |
| | | Chen GDP \| per sq km | | | | | x | x |
| *GEO* | Geographic & Land Conditions | NTL \| 1st lag | | | x | x | | x |
| | | NTL \| 2nd lag | | | x | x | | x |
| | | NTL \| growth | | | x | x | | x |
| | | NTL \| percentage increase | | | x | x | | x |
| | | NTL \| moving average 3 lags | | | x | x | | x |
| | | Electric consumption\| 1st lag | | | x | x | | x |
| | | Electric consumption\| 2nd lag | | | x | x | | x |
| | | Electric consumption \| moving average 3 lags | | | x | x | | x |
| | | Urban land \| percent | | | x | x | | x |
| | | Urban land \| area per sq km | | | x | x | | x |
| | | Crop land \| percent | | | x | x | | x |
| | | Crop land \|area per sq km | | | x | x | | x |
| | | ANP zone \| presence in NUTS 2 | | | x | x | | x |
| | | ANP zone \| count in NUTS 2 | | | x | x | | x |
| | | APN zone \| count per sq km | | | x | x | | x |
| | | Elevation differential (max-min) in NUTS2 | | | x | x | | x |
| | | Mountainous (elevation over 1600 m) | | | x | x | | x |
| | | Road \| count | | | x | x | | x |
| | | Road \| km2 | | | x | x | | x |
| | | Petrol \| count | | | x | x | | x |
| | | Petrol \| distance | | | x | x | | x |
| | | Petrol \| per sq km | | | x | x | | x |
| | | Legal mine \| presence | | | x | x | | x |
| | | Legal mine \| count | | | x | x | | x |
| | | Illegal mine \| presence | | | x | x | | x |
| | | Illegal mine \|count | | | x | x | | x |

| Group | Description | Variables | Model 1 (1) | Model 2 (2) | Model 3 (3) | Model 4 (4) | Model 5 (5) | Model 6 (6) |
|---|---|---|---|---|---|---|---|---|
| | | Hydroelectric plant \| count | | | x | x | | x |
| | | Hydroelectric plant \| per sq km | | | x | x | | x |
| | | Hydroelectric plant \| withing 50 km | | | x | x | | x |
| *INT* | Interactions | Population x Electric consumption | | | | x | | x |
| | | Population x NTL | | | | x | | x |
| | | Population x Urban land \| percent | | | | x | | x |
| | | Deforestation\| 1st lag Population | | | | x | | x |
| | | Deforestation\| 1st lag x Urban land \| percent | | | | x | | x |
| | | Deforestation x Crop land \| percent | | | | x | | x |
| | | NTL x Urban land \| percent | | | | x | | x |
| | | NTL Electric consumption | | | | x | | x |
| | | Road \| count x Urban land \| percent | | | | x | | x |
| | | Road \| km2 x Urban land \| percent | | | | x | | x |
| | | NTL^2 | | | | x | | x |
| | | Electric consumption^2 | | | | x | | x |
| Temporal Cross-Validation | | | | x | x | x | x | x |

/1 Neighboring variables were not computed because their processing requirements exceeded available resources.

**Table 5 Model Performance**

| Model | Complete Sample | | | Brazil Only Estimation | | |
|---|---|---|---|---|---|---|
| | Root Mean Squared Error | Median Average Error | $R^2$ | Root Mean Squared Error | Median Average Error | $R^2$ |
| Model 1 | 527.27 | 10.35 | 0.974 | 778.72 | 42.13 | 0.844 |
| Model 2 | 1313.06 | 47.7 | 0.931 | 2766.72 | 235.08 | 0.865 |
| Model 3 | 2555.48 | 25.04 | 0.975 | 2822.19 | 72.25 | 0.980 |
| Model 4 | 2625.10 | 26.98 | 0.971 | 3173.67 | 81.74 | 0.974 |
| Model 5 | 954.66 | 11.46 | 0.923 | 1256.36 | 28.35 | 0.953 |
| Model 6 | 881.76 | 11.6 | 0.957 | 1172.06 | 29.25 | 0.958 |

**Table 6. RMSE and $R^2$ results of Model 5 using Leave-One Country Out CV**

| Country | RMSE | $R^2$ |
|---|---|---|
| Brazil | 7264.6567 | 0.146 |
| Colombia | 2025.9521 | 0.372 |
| Peru (El) | 175.1133 | 0.412 |

## Table 7. RMSE of model 5 with Leave One NUTS1 Cross Validation

| Country | Region | 2007 | 2008 | 2009 | 2010 | 2011 | 2012 | 2013 | 2014 | 2015 | 2016 | 2017 | 2018 | 2019 | Average | $R^2$ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Brazil | Acre | 188.3 | 2,849.7 | 208.4 | 238.1 | 866.6 | 1,166.0 | 1,067.2 | 702.1 | 824.1 | 1,049.8 | 948.5 | 657.3 | 721.0 | 883.6 | 0.668 |
| | Amapa | 555.8 | 1,731.9 | 630.9 | 607.9 | 622.3 | 419.5 | 399.1 | 250.9 | 186.9 | 386.7 | 287.3 | 162.6 | 153.4 | 491.9 | 0.891 |
| | Amazonas | 3,998.4 | 3,979.6 | 3,999.4 | 4,867.3 | 2,686.3 | 2,710.3 | 2,711.3 | 2,039.2 | 712.9 | 1,379.4 | 523.3 | 2,015.8 | 2,314.1 | 2,610.6 | 0.798 |
| | Distrito Federal | 137,381.3 | 146,450.2 | 156,177.5 | 179,919.2 | 183,948.1 | 168,945.2 | 168,373.6 | 175,393.3 | 174,007.5 | 171,209.6 | 166,418.0 | 174,866.8 | 181,979.3 | 168,082.3 | - |
| | Goias | 264.2 | 255.7 | 350.7 | 547.6 | 566.7 | 1,020.9 | 655.5 | 638.8 | 1,809.4 | 658.1 | 475.1 | 425.3 | 884.2 | 657.9 | 0.756 |
| | Maranhao | 313.8 | 1,416.1 | 1,115.8 | 489.7 | 1,981.7 | 1,941.8 | 2,809.8 | 739.9 | 968.5 | 788.4 | 941.5 | 1,302.4 | 569.2 | 1,183.0 | 0.639 |
| | Mato Grosso | 337.6 | 419.0 | 406.4 | 377.2 | 1,398.2 | 853.8 | 783.7 | 645.1 | 660.0 | 602.6 | 637.2 | 766.7 | 606.8 | 653.4 | 0.900 |
| | Para | 1,316.5 | 2,290.1 | 1,950.3 | 2,574.4 | 3,492.8 | 3,012.5 | 3,038.7 | 3,334.1 | 3,235.9 | 3,010.1 | 3,398.3 | 3,135.1 | 3,628.0 | 2,878.2 | -0.078 |
| | Rondonia | 266.6 | 419.4 | 128.1 | 221.2 | 1,044.9 | 1,845.2 | 2,291.6 | 2,604.5 | 3,190.6 | 3,571.8 | 2,516.3 | 2,463.3 | 2,810.3 | 1,798.0 | -0.117 |
| | Roraima | 141.9 | 2,620.0 | 227.3 | 438.4 | 566.3 | 446.4 | 421.5 | 528.8 | 798.2 | 1,506.6 | 1,221.9 | 723.2 | 539.4 | 783.1 | 0.085 |
| | Tocantins | 150.9 | 214.5 | 237.2 | 202.6 | 469.0 | 756.2 | 597.7 | 562.5 | 479.7 | 507.2 | 503.9 | 386.6 | 300.2 | 412.9 | 0.582 |
| Colombia | Caqueta | | | | | | | 182.0 | 261.9 | 301.9 | 383.7 | 404.0 | 362.5 | 379.3 | 325.0 | 0.815 |
| | Cauca | | | | | | | 60.9 | 82.9 | 99.5 | 281.0 | 182.0 | 135.0 | 44.1 | 126.5 | -18.704 |
| | Guainia | | | | | | | 43.1 | 44.5 | 30.9 | 26.2 | 22.7 | 21.5 | 24.9 | 30.5 | 0.946 |
| | Guaviare | | | | | | | 52.1 | 34.3 | 31.4 | 81.9 | 19.3 | 133.7 | 37.9 | 55.8 | 0.963 |
| | Meta | | | | | | | 5,935.4 | 5,603.2 | 3,516.1 | 2,271.3 | 2,886.1 | 3,469.1 | 3,646.1 | 3,903.9 | 0.332 |
| | Narino | | | | | | | 144.6 | 448.3 | 841.6 | 1,619.2 | 1,436.4 | 1,436.4 | 1,600.2 | 1,075.2 | 0.898 |
| | Putumayo | | | | | | | 442.1 | 681.0 | 321.8 | 323.7 | 298.8 | 356.7 | 305.0 | 389.9 | 0.430 |
| | Vaupes | | | | | | | 50.8 | 36.2 | 32.8 | 44.1 | 53.5 | 54.8 | 38.6 | 44.4 | 0.832 |
| | Vichada | | | | | | | 167.5 | 160.7 | 273.2 | 367.0 | 290.7 | 199.5 | 174.7 | 233.3 | - |
| Peru | Ancash | 44.9 | 21.3 | 25.4 | 23.6 | 26.5 | 56.9 | 62.1 | 135.9 | 60.1 | 60.1 | 56.8 | 64.0 | | 53.1 | -6.456 |
| | Apurimac | 27.0 | 21.5 | 22.8 | 26.6 | 31.1 | 41.3 | 57.2 | 107.9 | 273.9 | 185.4 | 149.5 | 136.4 | | 90.1 | -0.363 |
| | Arequipa | 14.6 | 6.4 | 14.8 | 21.3 | 66.8 | 64.9 | 45.6 | 27.0 | 14.6 | 9.0 | 11.6 | 14.7 | | 25.9 | 0.528 |
| | Ayacucho | 341.6 | 114.6 | 170.1 | 462.1 | 191.4 | 356.0 | 424.5 | 96.5 | 273.7 | 106.8 | 106.0 | 121.2 | | 230.4 | -1.097 |
| | Cajamarca | 114.2 | 389.8 | 213.6 | 202.1 | 129.7 | 275.4 | 230.4 | 274.1 | 173.2 | 198.3 | 186.8 | 152.4 | | 211.7 | 0.560 |
| | Cusco | 77.1 | 76.6 | 89.8 | 114.5 | 137.4 | 244.3 | 181.4 | 174.5 | 150.0 | 132.9 | 130.1 | 128.7 | | 136.4 | 0.685 |

| Country | Region | 2007 | 2008 | 2009 | 2010 | 2011 | 2012 | 2013 | 2014 | 2015 | 2016 | 2017 | 2018 | 2019 | Average | $R^2$ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Huancavelica | 50.3 | 54.1 | 53.0 | 69.9 | 75.4 | 74.2 | 72.3 | 75.6 | 72.5 | 69.9 | 77.4 | 80.8 | | 68.8 | 0.481 |
| | Huanuco | 51.5 | 51.8 | 60.8 | 57.8 | 68.4 | 68.0 | 82.6 | 99.9 | 98.6 | 95.6 | 93.8 | 96.8 | | 77.1 | 0.708 |
| | Junin | 111.3 | 138.5 | 117.0 | 149.2 | 182.5 | 372.4 | 159.8 | 321.3 | 228.2 | 198.3 | 229.9 | 240.8 | | 204.1 | 0.642 |
| | La Libertad | 25.7 | 37.2 | 32.0 | 31.9 | 28.8 | 46.9 | 50.9 | 59.0 | 65.7 | 70.0 | 75.8 | 67.7 | | 49.3 | 0.547 |
| | Lambayeque | 7.2 | 9.9 | 19.1 | 15.9 | 7.0 | 8.7 | 10.2 | 11.2 | 13.2 | 12.4 | 16.3 | 15.8 | | 12.2 | - |
| | Loreto | 121.0 | 1,441.0 | 101.4 | 108.6 | 126.4 | 149.3 | 202.5 | 267.6 | 268.1 | 191.2 | 156.1 | 144.6 | | 273.1 | -2.748 |
| | Pasco | 148.0 | 127.0 | 129.0 | 104.9 | 103.2 | 164.1 | 191.7 | 212.1 | 183.8 | 175.1 | 180.3 | 166.8 | | 157.2 | -3.212 |
| | Piura | 5.1 | 16.1 | 10.5 | 11.0 | 4.0 | 6.3 | 5.5 | 4.7 | 7.4 | 7.3 | 7.8 | 7.3 | | 7.8 | -0.207 |
| | Puno | 10.6 | 10.9 | 9.5 | 9.1 | 19.4 | 27.0 | 27.1 | 29.3 | 30.9 | 31.4 | 34.6 | 36.5 | | 23.0 | -0.184 |
| | San Martin | 40.9 | 75.1 | 53.7 | 47.7 | 49.0 | 53.1 | 61.1 | 61.7 | 65.8 | 67.8 | 68.7 | 76.3 | | 60.1 | 0.762 |
| | Ucayali | 459.4 | 971.2 | 427.8 | 432.1 | 593.6 | 446.4 | 440.4 | 364.0 | 345.6 | 332.8 | 367.0 | 374.0 | | 462.9 | 0.493 |
| | Madre De Dios | 36.2 | 32.1 | 37.7 | 39.1 | 63.7 | 91.3 | 140.8 | 286.8 | 318.5 | 286.2 | 515.6 | 602.0 | | 204.2 | 0.143 |

*Note:* Missing RMSE values for Colombia and Peru reflect years in which official subnational GDP data were not available. A dash ("–") indicates NUTS1 units containing fewer than two municipalities, which makes computation of $R^2$ not feasible.

**Table 8. Moran's I of residuals by model, scope, weighting scheme, and year**

| Model | Scope | Weight Type | 2007 | 2008 | 2009 | 2010 | 2011 | 2012 | 2013 | 2014 | 2015 | 2016 | 2017 | 2018 | 2019 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | All countries | W | -0.500 | 0.181 | 0.067 | 0.968 | 1.635 | 2.004 | 2.432 | 1.848 | 1.709 | 9.725 | 11.474 | 4.672 | 3.357 |
| | | S | -0.293 | 0.310 | 0.227 | 1.046 | 1.755 | 2.115 | 2.664 | 2.050 | 1.927 | 10.478 | 11.732 | 4.835 | 3.302 |
| | | B | -0.103 | 0.419 | 0.362 | 1.087 | 1.823 | 2.164 | 2.855 | 2.248 | 2.139 | 11.066 | 11.884 | 4.986 | 3.275 |
| | | KNN | -0.497 | 0.250 | 0.154 | 1.192 | 1.797 | 2.081 | 2.650 | 1.350 | 1.491 | 8.739 | 11.897 | 4.227 | 3.989 |
| | Brazil | W | -0.557 | 0.094 | 0.048 | 0.201 | 0.497 | 0.507 | 0.860 | 0.822 | 0.407 | 1.372 | 1.026 | 1.084 | 1.375 |
| | | S | -0.386 | 0.185 | 0.140 | 0.281 | 0.583 | 0.605 | 0.967 | 0.934 | 0.593 | 1.539 | 1.233 | 1.238 | 1.504 |
| | | B | -0.229 | 0.260 | 0.213 | 0.342 | 0.643 | 0.673 | 1.046 | 1.009 | 0.742 | 1.661 | 1.400 | 1.348 | 1.590 |
| | | KNN | -0.538 | 0.205 | 0.169 | 0.410 | 0.698 | 0.661 | 1.011 | 0.841 | 0.458 | 1.297 | 1.259 | 1.216 | 1.885 |
| 2 | All countries | W | 2.981 | 2.251 | 2.018 | 2.171 | 2.743 | 2.620 | 3.207 | 3.623 | 3.218 | 2.608 | 2.734 | 2.568 | 1.081 |
| | | S | 2.724 | 2.077 | 1.891 | 2.016 | 2.544 | 2.478 | 3.024 | 3.293 | 2.940 | 2.366 | 2.502 | 2.410 | 0.995 |
| | | B | 2.433 | 1.874 | 1.736 | 1.830 | 2.302 | 2.292 | 2.798 | 2.969 | 2.660 | 2.143 | 2.277 | 2.239 | 0.894 |
| | | KNN | 2.872 | 2.181 | 1.994 | 2.066 | 2.737 | 2.697 | 3.077 | 3.303 | 2.971 | 2.375 | 2.533 | 2.452 | 1.161 |
| | Brazil | W | 2.725 | 1.502 | 1.023 | 1.016 | 1.265 | 1.279 | 1.570 | 1.663 | 1.471 | 1.132 | 1.077 | 1.078 | 1.027 |
| | | S | 2.304 | 1.304 | 0.915 | 0.914 | 1.134 | 1.156 | 1.417 | 1.469 | 1.309 | 1.021 | 0.984 | 1.006 | 0.962 |
| | | B | 1.887 | 1.101 | 0.797 | 0.799 | 0.986 | 1.018 | 1.246 | 1.263 | 1.135 | 0.896 | 0.877 | 0.914 | 0.874 |
| | | KNN | 2.496 | 1.468 | 1.088 | 1.064 | 1.357 | 1.405 | 1.663 | 1.752 | 1.554 | 1.255 | 1.183 | 1.213 | 1.157 |
| 3 | All countries | W | 2.914 | 2.139 | 1.847 | 2.153 | 2.784 | 2.404 | 3.778 | 3.923 | 3.562 | 2.959 | 2.689 | 2.448 | 0.935 |
| | | S | 2.647 | 1.985 | 1.719 | 1.972 | 2.601 | 2.276 | 3.581 | 3.634 | 3.255 | 2.675 | 2.453 | 2.315 | 0.873 |
| | | B | 2.355 | 1.803 | 1.567 | 1.763 | 2.378 | 2.115 | 3.336 | 3.332 | 2.939 | 2.410 | 2.229 | 2.168 | 0.797 |
| | | KNN | 2.803 | 2.082 | 1.805 | 1.987 | 2.855 | 2.434 | 3.481 | 3.545 | 3.176 | 2.621 | 2.437 | 2.329 | 1.024 |
| | Brazil | W | 1.986 | 1.364 | 0.787 | 0.998 | 0.988 | 1.162 | 1.795 | 1.765 | 1.820 | 1.257 | 1.178 | 1.104 | 0.967 |
| | | S | 1.686 | 1.188 | 0.708 | 0.889 | 0.906 | 1.043 | 1.612 | 1.567 | 1.604 | 1.125 | 1.058 | 1.014 | 0.906 |
| | | B | 1.382 | 1.007 | 0.620 | 0.768 | 0.807 | 0.912 | 1.412 | 1.358 | 1.380 | 0.981 | 0.925 | 0.911 | 0.827 |
| | | KNN | 1.941 | 1.360 | 0.875 | 1.018 | 1.164 | 1.322 | 1.867 | 1.831 | 1.789 | 1.328 | 1.248 | 1.208 | 1.092 |
| 4 | All countries | W | -0.455 | 38.816 | 0.251 | 0.435 | 0.402 | -0.804 | -0.136 | -2.463 | -2.734 | 2.591 | 2.866 | 0.297 | 3.471 |
| | | S | -0.255 | 40.954 | 0.290 | 0.444 | 0.414 | -0.562 | 0.166 | -2.001 | -2.577 | 2.611 | 2.895 | 0.274 | 2.998 |
| | | B | -0.066 | 42.529 | 0.354 | 0.452 | 0.442 | -0.299 | 0.551 | -1.495 | -2.368 | 2.591 | 2.822 | 0.486 | 2.679 |
| | | KNN | -0.331 | 31.066 | 0.195 | 0.578 | 0.407 | -0.594 | 0.292 | -2.142 | -3.142 | 2.605 | 3.489 | -0.762 | 4.393 |
| | Brazil | W | -1.083 | 22.405 | -0.225 | 0.487 | -0.053 | -0.474 | 1.185 | -3.051 | -2.531 | 1.601 | 1.637 | -0.540 | 3.386 |
| | | S | -0.821 | 23.597 | -0.126 | 0.456 | 0.071 | -0.300 | 1.350 | -2.421 | -2.409 | 1.535 | 1.563 | -0.363 | 3.346 |
| | | B | -0.569 | 24.464 | -0.011 | 0.431 | 0.167 | -0.160 | 1.521 | -1.778 | -2.239 | 1.444 | 1.446 | -0.139 | 3.226 |
| | | KNN | -0.913 | 17.848 | -0.156 | 0.575 | 0.129 | -0.363 | 1.748 | -3.094 | -2.734 | 1.851 | 1.886 | -0.817 | 5.233 |
| 5 | All countries | W | -0.966 | 20.807 | -0.989 | 0.134 | -0.121 | 3.640 | 1.706 | 0.228 | 0.877 | 5.708 | 4.335 | 2.622 | 3.163 |
| | | S | -0.778 | 22.122 | -0.768 | 0.253 | 0.034 | 3.040 | 1.676 | 1.180 | 1.795 | 5.726 | 4.392 | 3.022 | 3.054 |
| | | B | -0.584 | 23.144 | -0.550 | 0.350 | 0.161 | 2.476 | 1.716 | 2.205 | 2.715 | 5.757 | 4.385 | 3.659 | 2.988 |
| | | KNN | -0.776 | 15.921 | -0.785 | 0.519 | 0.489 | 3.764 | 1.424 | -1.242 | -0.654 | 4.519 | 4.365 | 0.611 | 4.467 |
| | Brazil | W | -2.710 | 5.387 | -0.824 | -0.369 | -0.096 | 2.903 | 0.235 | -1.391 | -1.565 | 1.853 | 0.909 | 0.447 | 2.093 |
| | | S | -2.178 | 5.992 | -0.658 | -0.224 | 0.031 | 2.459 | 0.326 | -0.577 | -0.889 | 1.976 | 0.990 | 0.864 | 2.119 |
| | | B | -1.663 | 6.508 | -0.493 | -0.096 | 0.143 | 2.034 | 0.436 | 0.235 | -0.256 | 2.071 | 1.028 | 1.353 | 2.112 |
| | | KNN | -2.164 | 4.043 | -0.636 | -0.050 | 0.340 | 3.053 | 0.504 | -1.742 | -1.814 | 1.640 | 1.270 | -0.925 | 3.563 |

*Note:* W = row-standardized weights (base); S = variance-stabilizing weights; B = binary contiguity weights; KNN = k-nearest neighbors weights. Significance at 10%, 5% and 1% would respectively be given if values are higher than 1.645, 1.96, 2.576.

**Table 9. Moran I statistics for predicted GDP estimates per model and scope**

| Model | Scope | 2007 | 2008 | 2009 | 2010 | 2011 | 2012 | 2013 | 2014 | 2015 | 2016 | 2017 | 2018 | 2019 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | All countries | 3.751 | 2.541 | 2.209 | 1.835 | 1.910 | 1.891 | 1.986 | 2.002 | 2.115 | 2.075 | 1.995 | 2.126 | 0.947 |
| | Brazil | 3.007 | 1.276 | 1.112 | 1.168 | 1.189 | 1.257 | 1.170 | 1.043 | 1.431 | 1.099 | 1.411 | 1.296 | 1.309 |
| 2 | All countries | 2.424 | 1.936 | 1.779 | 1.863 | 2.343 | 2.346 | 2.745 | 3.039 | 2.851 | 2.396 | 2.497 | 2.369 | 1.035 |
| | Brazil | 2.054 | 1.190 | 0.861 | 0.843 | 1.059 | 1.113 | 1.311 | 1.389 | 1.297 | 1.054 | 1.020 | 1.012 | 0.992 |
| 3 | All countries | 2.352 | 1.845 | 1.661 | 1.858 | 2.391 | 2.175 | 3.133 | 3.231 | 3.101 | 2.651 | 2.457 | 2.278 | 0.933 |
| | Brazil | 1.519 | 1.087 | 0.698 | 0.829 | 0.878 | 1.013 | 1.463 | 1.451 | 1.546 | 1.143 | 1.092 | 1.029 | 0.952 |
| 4 | All countries | 2.054 | 10.405 | 0.846 | 0.829 | 1.058 | 1.070 | 0.927 | 1.000 | 0.940 | 0.923 | 0.857 | 0.851 | 0.833 |
| | Brazil | 2.054 | 10.405 | 0.846 | 0.829 | 1.058 | 1.070 | 0.927 | 1.000 | 0.940 | 0.923 | 0.857 | 0.851 | 0.833 |
| 5 | All countries | 2.780 | 9.264 | 1.671 | 1.756 | 2.055 | 2.396 | 2.885 | 3.000 | 2.841 | 2.447 | 2.321 | 2.171 | 1.042 |
| | Brazil | 2.426 | 2.601 | 0.740 | 0.911 | 1.033 | 1.025 | 1.367 | 1.447 | 1.567 | 1.146 | 1.046 | 0.973 | 0.969 |

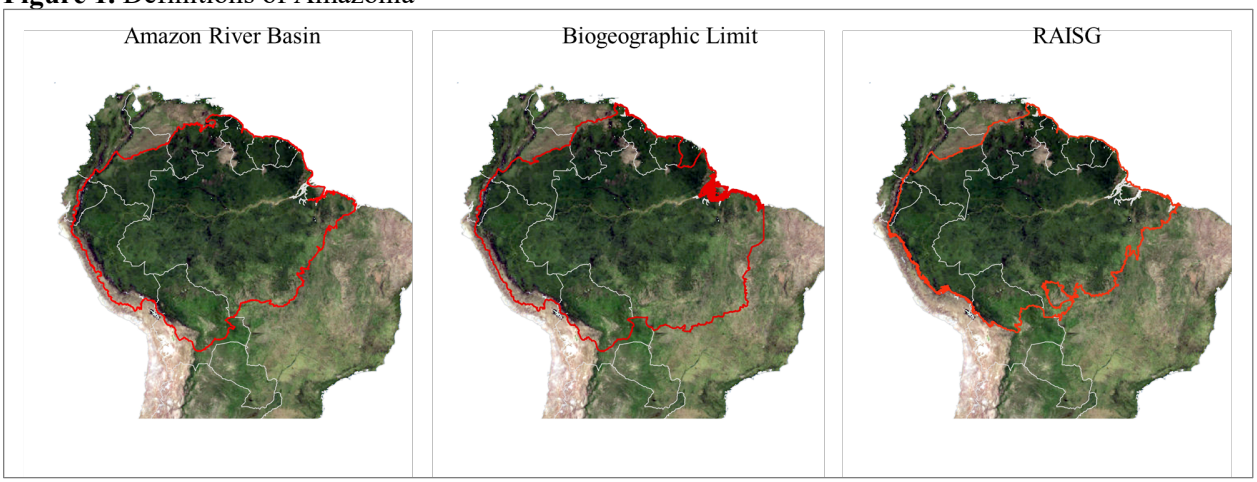
*Note*: Significance at 10%, 5% and 1% would respectively be given if values are higher than 1.645, 1.96, 2.576.

**Table 10. Official GDP Moran I's statistics per year**

| Model | 2006 | 2007 | 2008 | 2009 | 2010 | 2011 | 2012 | 2013 | 2014 | 2015 | 2016 | 2017 | 2018 | 2019 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| All Countries | 1.169 | 1.232 | 1.248 | 1.254 | 1.324 | 1.511 | 1.679 | 1.687 | 1.748 | 1.891 | 1.892 | 1.946 | 1.933 | 0.957 |
| Brazil | 0.472 | 0.492 | 0.480 | 0.492 | 0.540 | 0.640 | 0.725 | 0.736 | 0.768 | 0.849 | 0.849 | 0.878 | 0.860 | 0.932 |

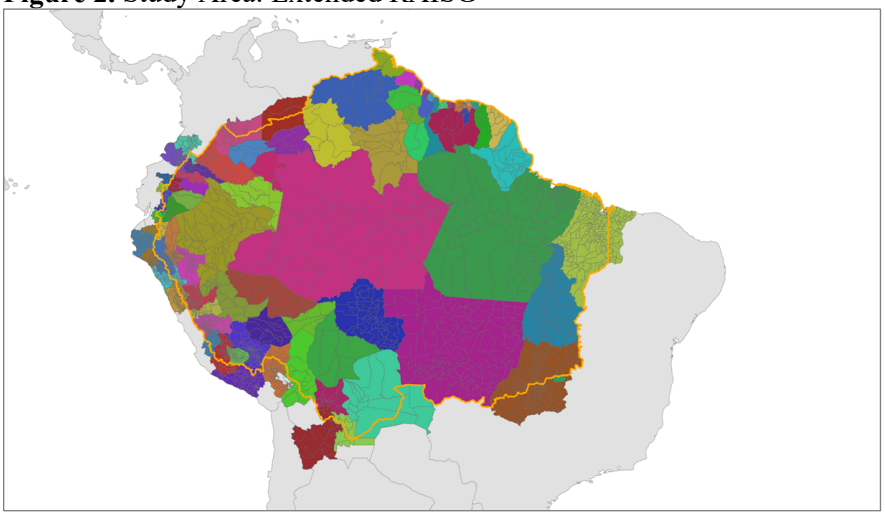
Note: Significance at 10%, 5% and 1% would respectively be given if values are higher than 1.645, 1.96, 2.576.

**Figure 1.** Definitions of Amazonia



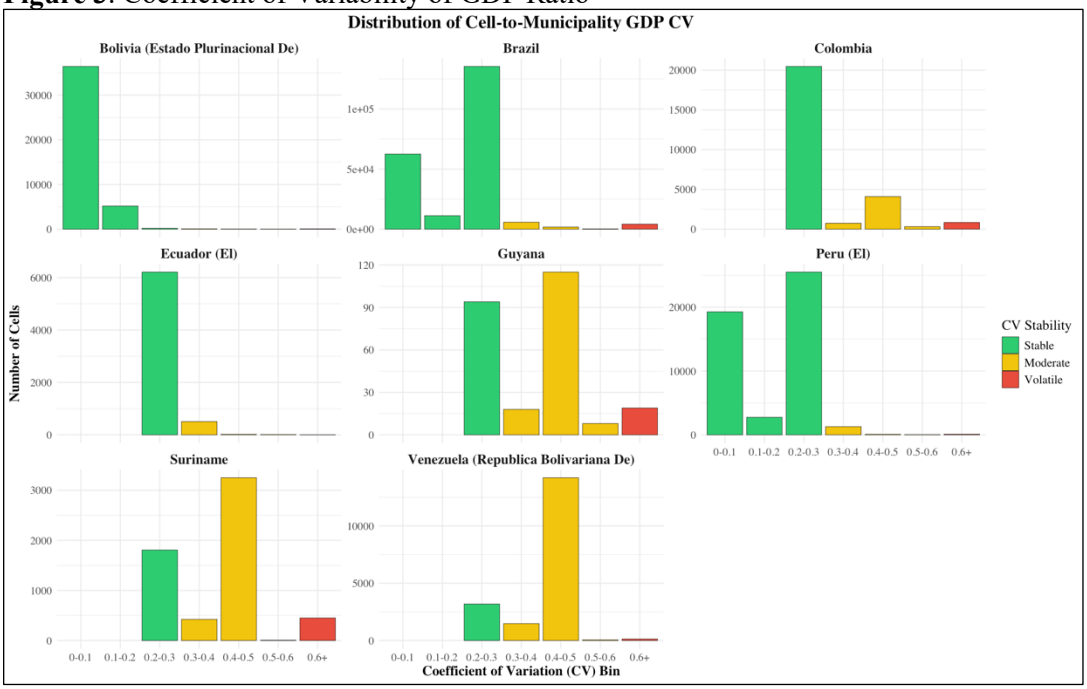

*Three primary definitions of Amazonia used in the literature, including the Amazon River Basin, biogeographic limits, and the RAISG delineation. The study adopts the RAISG definition as the baseline geographic framework (González Saldarriaga, et al., 2025).*

**Figure 2.** Study Area: Extended RAISG



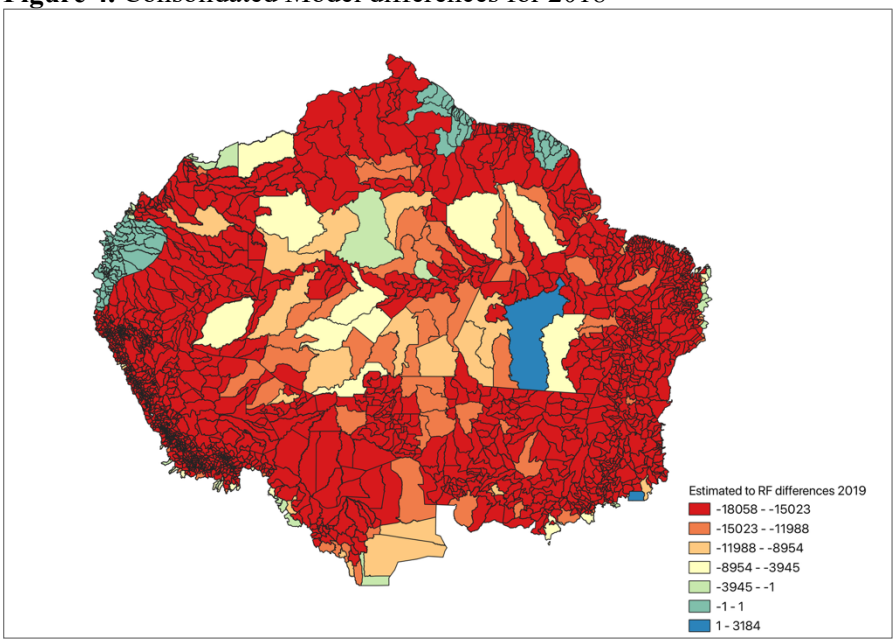
*Geographic extent of the study area, based on the RAISG definition, expanded to include complete subnational units intersecting the Amazonian boundary by at least 2%. This approach enables integration with official GDP statistics. Geographic extent of the study area was defined following the RAISG boundaries, expanded to include complete subnational units intersecting the Amazonian boundary by at least 2%. This criterion prevents artificial truncation of administrative areas while maintaining consistency with official GDP statistics. In applying this rule, we verified that major urban centers located outside the Amazon biome, such as Quito or Lima, were not erroneously included, whereas cities like La Paz and Brasília fall within the RAISG limits themselves and therefore remain part of the analytical scope. Although there is no standardized threshold for this type of spatial inclusion, our approach follows the reasoning discussed in Delafontaine et al. (2009), who highlight the need to address boundary effects in regional aggregation.*

**Figure 3**. Coefficient of Variability of GDP Ratio



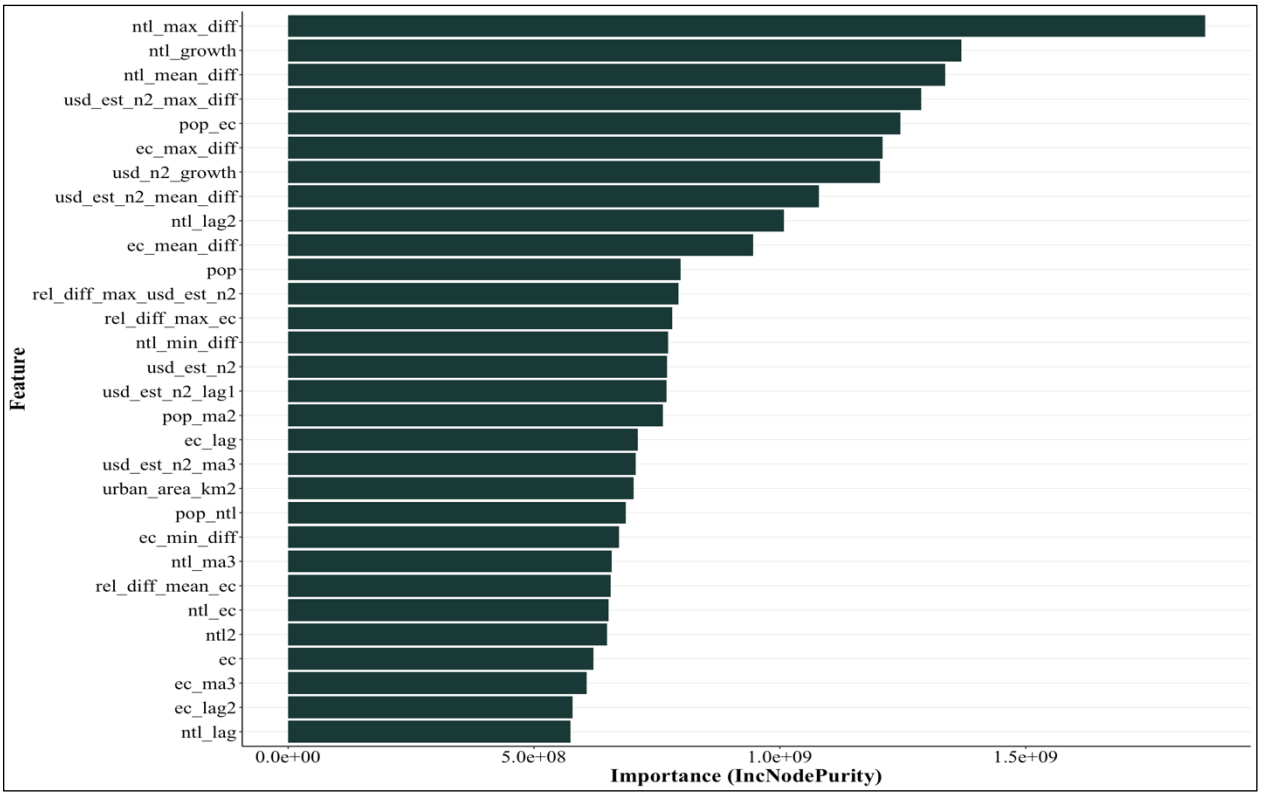

*Variation in GDP allocation ratios over time, showing greater stability in larger economies (e.g., Brazil and Peru) and higher volatility in smaller countries with less consistent official data.*
0

**Figure 4.** Consolidated Model differences for 2018



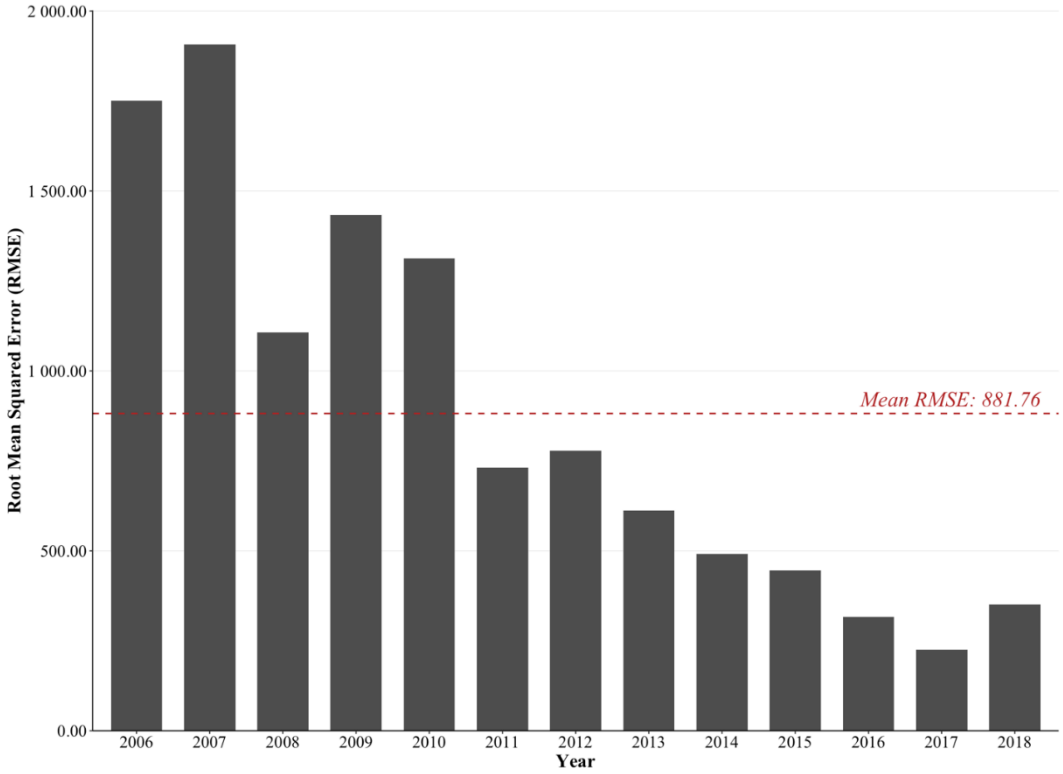

*Spatial distribution of GDP estimation differences between the consolidated model and the Chen et al. (2021) NTL-based GDP dataset for the year 2018, at the NUTS2 level. Blue regions indicate areas where the consolidated model estimates higher GDP, while red indicates lower estimates.*

**Figure 5.** Feature Importance for the Full Model (Model 5, Complete Sample)




*Relative importance of predictor variables in Model 5, as determined by the Random Forest algorithm. The maximum Nighttime light difference between the unit and its neighbors is the most important feature for the model. Also, differences in estimation of GDP as well as the interaction of population and Electricity consumption are important. For variable description, see Table.*

**Figure 6.** Consolidated Model's RMSE per year.




*Root Mean Squared Error (RMSE) of Model 5 (complete sample) predictions over time. Lower RMSE in more recent years reflects the increased availability of training data and improved model stability./****