

# Cuadernos de Desarrollo Económico

83

## Monitoreo Trimestral de la Actividad Económica en Bogotá con Datos Satelitales y Movilidad

FEBRERO 2026

Eduard Fernando Martínez-González  
Miguel Angel Torres Montoya







Carlos Fernando Galán Pachón  
**Alcalde Mayor de Bogotá**

**Secretaría de Desarrollo Económico**

María del Pilar López Uribe  
**Secretaria de Desarrollo Económico**

José Fuentes Ortega  
**Subsecretario de Desarrollo Económico**

Gabriel Hernando Angarita Tovar  
**Director de Estudios de Desarrollo Económico**

Pilar Torres Alvarado  
**Subdirectora de Estudios Estratégicos**

Yaneth Lucia Pinilla Beltrán  
**Subdirectora de Información y Estadísticas**

**Autores**

Eduard Fernando Martínez-González  
Secretaría de Desarrollo Económico de Bogotá y Centro de  
Investigación en Economía y Finanzas (CIENFI)  
Miguel Angel Torres Montoya  
Centro de Investigación en Economía y Finanzas (CIENFI)

**Material fotográfico**

Archivo SDDE y Alcaldía de Bogotá

Cuaderno No. 83  
Monitoreo Trimestral de la Actividad Económica  
en Bogotá con Datos Satelitales y Movilidad

**Revisión de estilo**

Pilar Torres Alvarado

**Diagramación**

César Javier Daza Gutiérrez

**ISSN**

2981-4790 (En línea)

Puede encontrar información editorial de este cuaderno en la página Web:

<http://observatorio.desarrolloeconomico.gov.co>

Contáctenos en:

Secretaría Distrital de Desarrollo Económico

Calle 28 No 13 A- 35, Edificio Centro de comercio Internacional. Bogotá, D.C.

Teléfono: 601 369 3777 Ext: 204

[observatorio@desarrolloeconomico.gov.co](mailto:observatorio@desarrolloeconomico.gov.co)



La serie de Cuadernos de Desarrollo Económico es una publicación de la Secretaría de Desarrollo Económico de la Alcaldía Mayor de Bogotá. Los trabajos de la Serie de cuadernos son de carácter provisional; Las opiniones y el contenido son responsabilidad exclusiva de los autores y no comprometen a la Secretaría de Desarrollo Económico ni a la Alcaldía Mayor de Bogotá. Todo el material está protegido por los derechos de autor y su uso está permitido libremente en tanto, se haga mención y se cite en la bibliografía. Su versión en línea está disponible bajo la licencia **Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International**

# Monitoreo Trimestral de la Actividad Económica en Bogotá con Datos Satelitales y Movilidad

## RESUMEN

Este documento presenta un sistema de monitoreo trimestral de la actividad económica en Bogotá a partir de señales satelitales de luminosidad nocturna, registros operativos de movilidad y estructura demográfica urbana. Se construye un panel trimestral 2012Q2–2025 a nivel de grilla de 450×450 m que integra radiancia VIIRS-DNB (con control explícito de nubosidad), validaciones y salidas del sistema TransMilenio, y población y viviendas del Censo 2018. El PIB oficial de la ciudad se espacializa mediante un *pseudo-target* que distribuye el PIB per cápita entre celdas según la población residente, garantizando coherencia contable con las cuentas nacionales. Sobre esta base se entrena un modelo de red neuronal feedforward, junto con especificaciones alternativas de gradient boosting y bosques aleatorios, calibrados con validación temporal estricta y un barrido extensivo de hiperparámetros. Los modelos seleccionados reproducen el PIB agregado con errores inferiores al 1 % del nivel medio y tasas de crecimiento para 2024 muy próximas a las cifras oficiales. El sistema genera predicciones trimestrales de PIB por grilla para 2025 basadas exclusivamente en señales de alta frecuencia, lo que permite caracterizar la dinámica económica intraurbana y proporciona insumos para tableros de monitoreo territorial, análisis de crecimiento formal e informal y diseño de políticas de desarrollo económico local.

**Palabras clave:** actividad económica intraurbana; luces nocturnas VIIRS; movilidad urbana; redes neuronales; PIB subnacional; datos de alta frecuencia; Bogotá.



# ABSTRACT

This paper develops a quarterly monitoring system of economic activity in Bogotá using satellite night lights, operational mobility records and urban demographic structure. We build a quarterly panel for 2012Q2–2025 at a 450×450 m grid level that combines VIIRS-DNB radiance (with an explicit cloud-free control), TransMilenio validations and exits, and population and housing from the 2018 Population Census. Official city GDP is spatialized through a *pseudo-target* that allocates GDP per capita across grid cells according to resident population, ensuring consistency with national accounts. On top of this prior, we train a feedforward neural network, along with gradient boosting and random forest specifications, calibrated under strict temporal validation and an extensive hyperparameter search. The selected models reproduce aggregate GDP with errors below 1 % of the mean level and deliver growth rates for 2024 that closely match official figures. The system then generates grid-level quarterly GDP predictions for 2025 based solely on high-frequency external signals, allowing us to characterize intra-urban economic dynamics and to feed territorial monitoring dashboards, analyses of formal–informal growth and the design of local economic development policies.

**Keywords:** intra-urban economic activity; VIIRS night lights; urban mobility; neural networks; subnational GDP; high-frequency data; Bogotá.

# CONTENIDO

<b>1</b>	<b>Introducción</b>	<b>8</b>
----------	---------------------	----------

---

<b>2</b>	<b>Datos</b>	<b>11</b>
	2.1. Fuentes y arquitectura	11
	2.2. Armonización de los conjuntos de datos	12
	2.3. Descriptivas	13

---

<b>3</b>	<b>Metodología</b>	<b>21</b>
	(i) Prior y <i>pseudo-target</i>	21
	(ii) Modelo, features y entrenamiento	22
	(iii) Validación y selección de modelo	23

---

<b>4</b>	<b>Resultados</b>	<b>24</b>
	4.1. Espacio de búsqueda	25
	4.2. Desempeño del modelo	27

---

<b>5</b>	<b>5. Conclusiones</b>	<b>29</b>
----------	------------------------	-----------

---

<b>6</b>	<b>6. Bibliografía</b>	<b>30</b>
----------	------------------------	-----------





# 1. INTRODUCCIÓN

---

La medición de la actividad económica a escalas intraurbanas ha avanzado de forma sustantiva gracias a la disponibilidad creciente de señales de alta frecuencia provenientes de teledetección y trazas digitales. Los trabajos fundacionales de Doll, Muller y Morley (2006) establecieron el uso de imágenes nocturnas como aproximación a la actividad subnacional. Investigaciones más recientes han extendido este enfoque empleando productos satelitales de mayor resolución (Coetzee et al., 2021) y algoritmos de aprendizaje automático aplicados sobre imágenes diurnas (Lehnert et al., 2023), lo que ha permitido mejorar la representación espacial y temporal de la dinámica económica urbana. De forma paralela, los registros operativos de movilidad han demostrado ser trazadores robustos de intensidad funcional en sistemas de transporte masivo, ofreciendo información granular sobre flujos laborales y comerciales dentro de las ciudades (Kreindler y Miyauchi, 2023).

Recientemente la literatura sobre uso de luces nocturnas como aproximación a la actividad económica, ha mostrado una larga tradición dentro de la economía espacial y el análisis regional. El estudio de Doll et al. (2006) demostró una correlación robusta entre la extensión de áreas iluminadas y diversos parámetros socioeconómicos a nivel nacional y subnacional. Esta aproximación fue posteriormente formalizada por Henderson et al. (2012), quienes desarrollaron un marco estadístico para utilizar el crecimiento de la luminosidad como un indicador del crecimiento real y como un complemento a las cuentas nacionales. De manera similar, Chen y Nordhaus (2011) mostraron que las luces nocturnas permiten reconstruir patrones de producción en países con sistemas estadísticos incompletos o con ausencia de censos recientes.

A pesar de esta evidencia agregada, varios estudios han documentado que la relación entre luminosidad y actividad económica no es completamente estable a escalas subnacionales. Bickenbach et al. (2016) encuentran que la elasticidad luz-PIB varía sustancialmente entre regiones dentro de países como Brasil, Estados Unidos, India y varias economías europeas, lo que sugiere que la señal luminosa captura de forma heterogénea las estructuras económicas locales. De manera consistente, Lessmann y Seidel (2016) muestran que la capacidad predictiva de la luminosidad depende críticamente de factores como el nivel de desarrollo, la densidad poblacional y la composición sectorial, indicando que el uso exclusivo de luces nocturnas puede resultar insuficiente para caracterizar economías territoriales complejas.

Frente a estas limitaciones, avances recientes han integrado datos satelitales con métodos de aprendizaje automático y variables auxiliares, mejorando la precisión de las estimaciones subnacionales. Estudios para Bangladesh (Ehsan et al. 2020), Egipto (Suleiman 2021) y Paraguay (McCord y Rodríguez-Heredia 2022) muestran que la combinación de luces nocturnas VIIRS-DNB con datos de densidad poblacional incrementa de manera importante la capacidad predictiva del PIB a nivel regional. De forma complementaria, Liu et al. (2021) utilizan imágenes diurnas de alta resolución y una red neuronal convolucional para extraer características del entorno construido que enriquecen la información nocturna y permiten obtener estimaciones más finas de la actividad económica local.

Finalmente, la alta informalidad característica de las ciudades latinoamericanas reduce la cobertura y consistencia de los registros administrativos, mientras que los agregados oficiales se publican con rezagos y en unidades espaciales extensas que enmascaran la heterogeneidad interna. La ausencia de métricas oportunas y comparables limita el diseño de políticas focalizadas y dificulta el seguimiento de la dinámica económica en territorios con fuerte presencia de informalidad. Este contexto exige metodologías capaces de combinar señales externas al sistema estadístico tradicional con estructuras demográficas y funcionales urbanas.

Este estudio aborda esta brecha mediante un esquema de estimación del PIB a nivel de grilla que integra información satelital, movilidad del transporte masivo y estructura poblacional, evitando el uso de registros administrativos en la etapa predictiva. La metodología produce una serie trimestral de *PIB por grilla* de 450×450 m para las cinco ciudades más grandes de Colombia, aprovechando señales exógenas de alta frecuencia y un *pseudo-target* anclado en el PIB per cápita oficial. Esta aproximación permite capturar variación espacial y temporal sin depender de fuentes administrativas sujetas a cobertura diferencial, discontinuidades o cambios regulatorios.

El proceso se estructura en dos etapas. Primero, se construye un *pseudo-target* que distribuye el PIB per cápita de cada ciudad entre las celdas de acuerdo con la población del Censo 2018, tratada como invariante en el tiempo. Esta capa establece una línea base coherente con los agregados oficiales y define el nivel del PIB por celda. Segundo, se entrena un modelo predictivo basado en una red neuronal *feedforward* que mapea radiancia VIIRS-DNB (con control explícito de nubosidad), flujos y estacionalidad del sistema TransMilenio, y estructura demográfica hacia una estimación refinada del PIB a nivel de grilla. El modelo utiliza ponderaciones basadas en confiabilidad y validación temporal tipo *rolling origin* para preservar la correcta estructura cronológica de la serie.



El entrenamiento se realiza con datos trimestrales entre 2012Q2 y 2023Q4, siguiendo una partición temporal estricta tipo 70–30 % para evitar fugas de información. La validación externa se ejecuta mediante la comparación entre las predicciones agregadas y las proyecciones oficiales del PIB de 2024, permitiendo evaluar desempeño fuera de muestra en un contexto de variación macroeconómica reciente. Posteriormente, el modelo seleccionado genera pre-dicciones trimestrales para 2025, construidas exclusivamente sobre señales de alta frecuencia y la estructura aprendida durante el período histórico.

La contribución del estudio es tres aspectos. Primero, se produce una serie trimestral de alta resolución espacial (450×450 m) que captura la estructura económica urbana sin depender de registros administrativos en el proceso de estimación. Esta serie permite analizar la dinámica económica intraurbana con granularidad consistente en el tiempo y replicable entre ciudades. Segundo, se desarrolla una metodología de descomposición formal–informal que emplea información administrativa únicamente en una etapa pospredictiva, evitando dependencia mecánica entre predictores y resultados. El marco es escalable a otras ciudades de América Latina y el Caribe y permite generar herramientas operativas — mapas trimestrales, table-ros abiertos y reportes ejecutivos— que facilitan el seguimiento territorial de la actividad económica, la planeación urbana y la formulación de políticas de inclusión productiva.

Finalmente, este documento se estructura de la siguiente manera. La Sección 2 describe las fuentes de información, los procedimientos de armonización espacial y temporal, y las principales características del panel de datos utilizado. La Sección 3 presenta la metodología de estimación, incluyendo la construcción del *pseudo-target*, la arquitectura del modelo predictivo y el esquema de validación temporal. La Sección 4 expone los resultados del entrenamiento, el desempeño de los modelos y las predicciones agregadas y espaciales del PIB. La Sección 5 ofrece las conclusiones y plantea la agenda de trabajo futura, centrada en la desagregación territorial del crecimiento y en el uso de señales de alta frecuencia para el monitoreo económico en tiempo casi real.

## 2. DATOS

---

Esta sección presenta las fuentes utilizadas para estimar el PIB por el enfoque de demanda a nivel de grilla, los procedimientos de integración espacial y temporal, y las características principales de los conjuntos de datos empleados.

### 2.1. FUENTES Y ARQUITECTURA

La construcción del conjunto de datos integra fuentes con distinta granularidad, frecuencia y procesos de generación, todas mapeadas a una unidad espacial de análisis común con periodicidad trimestral. La primera fuente corresponde a las imágenes satelitales de radiancia nocturna VIIRS-DNB<sup>1</sup>, un sensor que captura radiancia en el rango visible–infrarrojo cercano con sensibilidad suficiente para detectar emisiones lumínicas de baja intensidad, incluyendo alumbrado público, iluminación residencial dispersa e infraestructura comercial de mediana escala. Su resolución radiométrica permite identificar variaciones marginales asociadas a cambios en la actividad urbana.

VIIRS-DNB registra información nocturna mediante los satélites Suomi-NPP y NOAA-20, ambos en órbita polar sincrónica con el Sol, que realizan un sobrevuelo nocturno por día sobre Colombia, generando hasta dos oportunidades de observación por noche en las zonas cubiertas por ambas plataformas. En este documento se emplea el compuesto mensual VNL VCMCFG desarrollado por el Earth Observation Group (EOG), el cual aplica un procedimiento estandarizado de depuración que elimina nubosidad y fuentes temporales de luz. Este compuesto produce una medición mensual de radiancia agregada a una resolución de 450×450 m por celda, a partir de la cual se construyen métricas trimestrales de radiancia promedio y número de días despejados para cada grilla de 450×450 m.

La segunda fuente corresponde a los registros operativos de movilidad del sistema TransMilenio. Estos datos consisten en conteos diarios de entradas y salidas validados por torniquete para cada estación y troncal del sistema, con georreferenciación precisa a nivel de infraestructura. A partir de esta base diaria se construyen agregaciones operativas trimestrales por estación o troncal, desagregadas por franjas horarias, que permiten capturar movimientos laborales y comerciales con alta resolución espacial y temporal.

---

<sup>1</sup> VIIRS-DNB (*Visible Infrared Imaging Radiometer Suite – Day/Night Band*) es un sensor de observación nocturna operado por la National Oceanic and Atmospheric Administration.



La tercera fuente corresponde a la población residente del Censo Nacional de Población de 2018 del DANE, asignada a cada celda mediante intersección exacta de manzanas censales. Esta capa se mantiene invariante durante el horizonte analizado y se utiliza para anclar el *pseudo-target* inicial del PIB por grilla, además de operar como control estructural de densidad urbana. Adicionalmente, del DANE se emplean las series oficiales del PIB trimestral de Bogotá a precios constantes, que sirven como agregados de referencia para la construcción del PIB per cápita y para la validación externa del modelo. Finalmente, los registros empresariales provienen del Registro Único Empresarial y Social (RUES) y se utilizan exclusivamente como insumos pospredictivos dentro de la etapa de descomposición del crecimiento. Estos registros aportan información sobre la localización y magnitud del tejido empresarial formal, pero no participan en ninguna fase del entrenamiento del modelo predictivo.

## 2.2. ARMONIZACIÓN DE LOS CONJUNTOS DE DATOS

La construcción del panel final requiere integrar fuentes heterogéneas en una unidad espacial común y en una frecuencia temporal trimestral. El producto resultante es un panel balanceado grilla-trimestre para las 2,100 celdas de 450×450,m que cubren la superficie urbana consolidada de Bogotá. Cada celda se observa reiteradamente a lo largo del horizonte 2012Q2–2025, con un vector de características invariante o trimestral según la naturaleza de la fuente: población y viviendas (Censo 2018), radiancia y días despejados (VIIRS-DNB), flujos operativos de movilidad (TransMilenio) y conteos empresariales (RUES).

El proceso de armonización inicia definiendo la grilla como unidad de observación, dado que constituye el único soporte espacial capaz de recibir todas las fuentes mediante operaciones de intersección o asignación espacial. La población residente y el número de viviendas se integran mediante una intersección exacta entre las manzanas censales del DANE y las celdas de la grilla: para cada celda se calcula el número de personas y de viviendas contenidas dentro de su perímetro. De igual forma, los registros empresariales del RUES —ya geocodificados a nivel de punto— se asignan a las celdas mediante conteo directo del número de establecimientos ubicados en cada unidad espacial.

La asignación espacial de los datos de movilidad se realiza en dos etapas. Primero, cada corredor troncal se intersecta con la grilla regular de 450×450,m, de modo que las celdas que contienen infraestructura troncal reciben directamente los conteos diarios correspondientes a esa troncal. Segundo, para las celdas que no intersectan ninguna troncal se implementa un esquema de asignación mediante ponderación inversa a la distancia: cada celda recibe una combinación convexa de los flujos diarios observados en las troncales más cercanas, donde el peso asignado a cada troncal decrece con la distancia entre el centroide de la celda y el eje troncal. Finalmente, la radiancia nocturna y los días despejados —ya provistos en formato raster— se integran directamente mediante extracción zonal del valor promedio y del conteo de observaciones válidas en cada celda para cada trimestre.

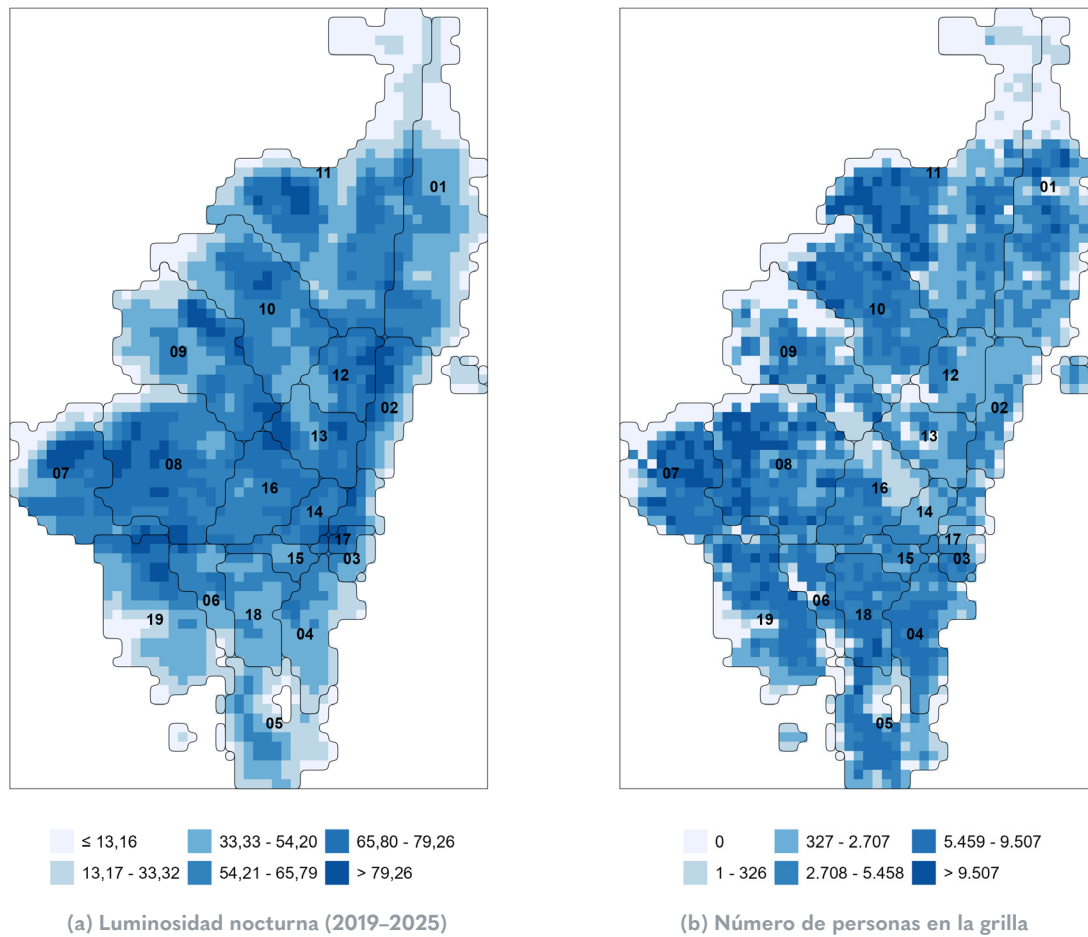
## 2.3. DESCRIPTIVAS

Las descriptivas del conjunto de datos a nivel de grilla se organizan en cuatro componentes: (i) distribución espacial, (ii) distribuciones marginales, (iii) correlaciones bivariadas y (iv) comovimientos temporales entre las variables del modelo y el PIB agregado. Las Figuras 1 y 2 presentan la distribución espacial de las principales variables empleadas en la estimación, para un universo de 2.100 celdas de 450×450,m que cubren la superficie urbana consolidada de Bogotá.

La Figura 1 presenta la distribución espacial del valor promedio de radiancia VIIRS-DNB para el período 2019–2025 y de la población censal de 2018. Las grillas se clasifican en seis intervalos definidos a partir de los cuantiles empíricos (10, 25, 50, 75 y 95). Se identifican tres hechos centrales. Primero, existen áreas donde la luminosidad y la densidad poblacional coinciden, particularmente en sectores del suroccidente (Comuna 11) y en las comunas 7 y 8, lo que refleja zonas residenciales consolidadas con actividad urbana sostenida. Segundo, algunos píxeles del nororiente —específicamente en la Comuna 1— exhiben una correspondencia más estrecha entre densidad poblacional baja o moderada y niveles de radiancia igualmente bajos o moderados, lo que sugiere una estructura predominantemente residencial con menor intensidad funcional nocturna.

Tercero, una fracción amplia del territorio presenta niveles altos de radiancia sin correspondencia proporcional en población. En la Comuna 9, los píxeles del aeropuerto se ubican por encima del percentil 95 de radiancia con población prácticamente nula. Un patrón similar se observa en la franja que conecta de occidente a oriente las comunas 9, 16 y 13, donde la luminosidad elevada obedece a infraestructura vial y actividad logística —en particular el corredor de la Calle 26— sin implicar concentración residencial. De igual forma, en la zona nororiental (Comuna 2), los niveles de radiancia se mantienen altos pese a densidades poblacionales moderadas, en coherencia con la presencia de corredores comerciales y de servicios. Estos hechos confirman que la relación entre luminosidad nocturna y densidad residencial es no lineal y está mediada por la función urbana predominante en cada zona.

**Figura 1: Distribución espacial de luminosidad nocturna y población en Bogotá**



**Nota:** La luminosidad corresponde al promedio trimestral de radiancia VIIRS-DNB (producto VCMCFG) para 2019–2025. La población proviene del Censo 2018 y se asigna a la grilla de 450×450 m mediante intersección espacial.

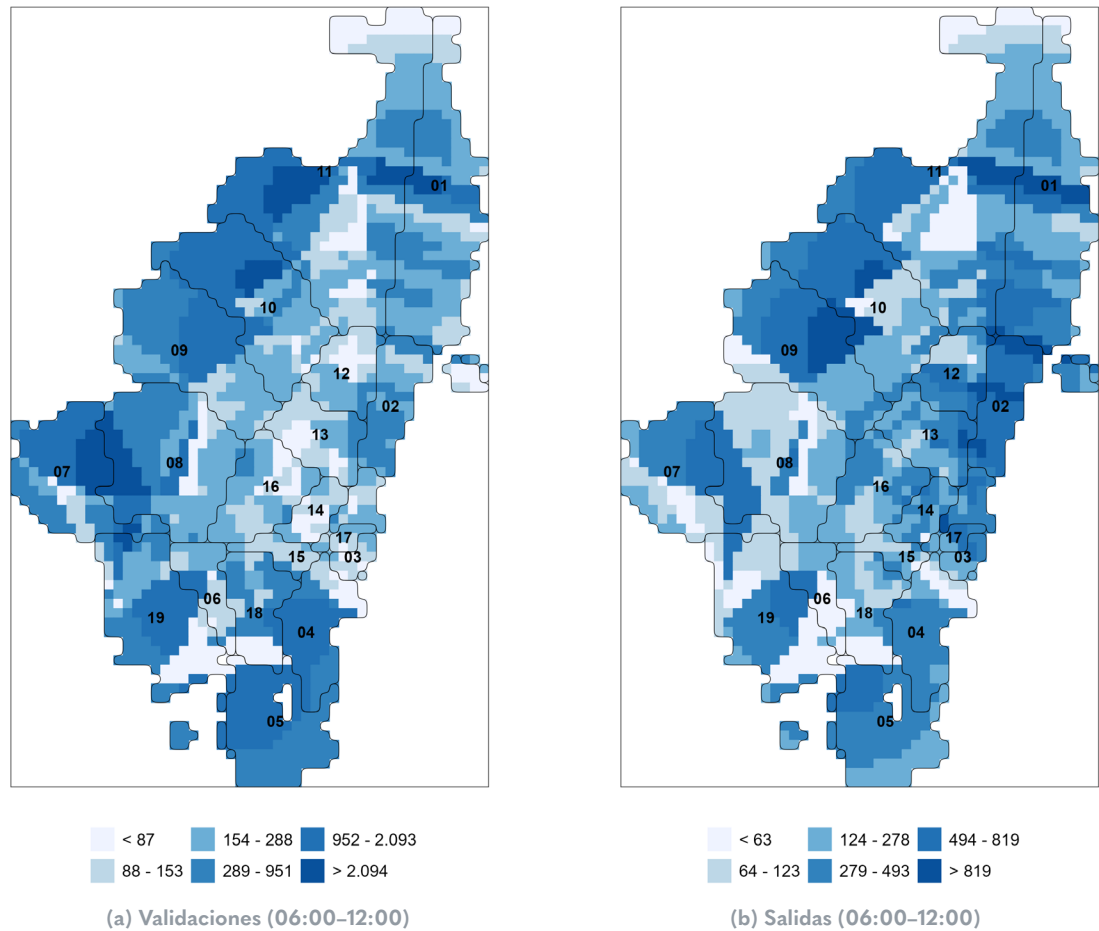
**Fuente:**

Elaboración: Secretaría Distrital de Desarrollo Económico - Observatorio de Desarrollo Económico de Bogotá.

La Figura 2 presenta la distribución espacial del promedio de validaciones y salidas registradas en las estaciones y troncales de TransMilenio entre las 06:00 y las 12:00 para el período 2019–2025. Las grillas se clasifican en seis intervalos definidos por los cuantiles 10, 25, 50, 75 y 95 de cada variable. La figura evidencia patrones diferenciados entre los puntos de origen y destino de los desplazamientos en horas de la mañana. Las validaciones se concentran en el occidente, particularmente en sectores residenciales de las comunas 7, 8 y 19, configurando zonas emisoras de viajes. En contraste, las salidas se agrupan en el oriente, en especial en la Comuna 2, donde se localiza la mayor parte del centro empresarial, institucional y financiero de Bogotá.

El mapa también revela comportamientos atípicos respecto al patrón general. En la Comuna 9, los niveles de salidas son elevados pese a presentar volúmenes moderados de validaciones, lo que es consistente con la presencia de infraestructura logística y nodos de transferencia próximos al aeropuerto. En el extremo norte, las estaciones cercanas al Portal del Norte concentran un volumen elevado de salidas. Finalmente, en la Comuna 2 se observa una combinación de salidas altas y validaciones bajas; esta asimetría indica que, aunque existe presencia residencial, una proporción sustantiva de los desplazamientos hacia esta zona proviene de otras comunas, coherente con su rol como principal destino laboral y de servicios.

**Figura 2: Validaciones y salidas por troncales de TransMilenio**



**Nota:** Las validaciones y las salidas corresponden a registros del sistema TransMilenio agregados por troncales. Las cifras se expresan en miles de personas y representan promedios entre semana en el intervalo 06:00–12:00.

Fuente:

Elaboración: Secretaría Distrital de Desarrollo Económico - Observatorio de Desarrollo Económico de Bogotá.

La Tabla 1 resume las propiedades estadísticas del vector de características  $X_{gt}$  para el universo de 2,100 celdas trimestrales. El PIB por grilla exhibe una distribución altamente asimétrica, con una mediana de 23,008 millones de COP constantes y un percentil 95 superior a 82,000 millones, lo que evidencia una concentración marcada de la actividad económica en un subconjunto reducido de celdas. Esta heterogeneidad es consistente con la forma en que el nivel del PIB se distribuye espacialmente a partir de la densidad poblacional, la cual opera como ancla para el *pseudo-target*. La estructura de concentración y dispersión de esta variable será examinada más adelante al analizar su co-movimiento con las señales de movilidad y luminosidad.

La radiancia VIIRS presenta una media de 49.3 unidades y un rango intercuartílico entre 31.7 y 66.3 unidades, reflejando diferencias sustantivas en intensidad lumínica entre sectores urbanos. Los días despejados muestran variación amplia ( $P25 = 6$ ;  $P75 = 15$ ), condición relevante para la calidad radiométrica de la señal. Este indicador funciona como un control explícito de la confiabilidad del dato satelital: el promedio trimestral de radiancia se construye ponderando cada observación mensual por el número de veces que el píxel es captado sin nubosidad en el compuesto VCMCFG. En consecuencia, las celdas con mayor frecuencia de observaciones libres de nubosidad aportan información más robusta al estimador final de luminosidad, mientras que aquellas con menor claridad atmosférica reciben un peso relativo menor en la agregación trimestral.

Por otra parte, las variables demográficas muestran que la población residente alcanza una media de 3,385 personas por celda, con una dispersión elevada (desviación estándar de 3,200) y un tramo superior que supera las 9,500 personas. El número de viviendas presenta un patrón similar, con una mediana de 1,042 unidades y un percentil 95 cercano a 3,200, lo que confirma la distribución fuertemente concentrada de la densidad habitacional observada en la Figura 1. La actividad empresarial, medida por el conteo de firmas registradas, presenta una media de 134 unidades y un percentil 95 de 423, indicando una cola superior más pronunciada y consistente con la distribución espacial de zonas comerciales e industriales.

Las variables de movilidad presentan patrones diferenciados entre franjas horarias y entre tipos de registro. Aunque los promedios de salidas y validaciones son de magnitudes similares, la distribución revela asimetrías relevantes. Entre las 00:00 y las 06:00, las validaciones superan de manera sistemática a las salidas, lo que indica que una proporción sustantiva de los viajes se inicia en horas de la madrugada y se materializa en desplazamientos efectivos durante el intervalo 06:00–12:00, franja en la que se concentran los mayores volúmenes de salidas. Además, las validaciones exhiben percentiles 95 considerablemente más altos que las salidas, especialmente en 06:00–12:00, donde alcanzan más de 2.3 millones de registros por trimestre.



Este patrón implica la existencia de un número reducido de estaciones que funcionan como nodos emisores de gran escala, concentrando una fracción desproporcionada de los inicios de viaje en la ciudad. En contraste, los percentiles superiores de salidas no presentan magnitudes comparables, lo que sugiere que dichos nodos emisores dispersan pasajeros hacia una red amplia de destinos en lugar de concentrar la llegada en puntos equivalentes de gran tamaño. La estructura conjunta de medias y percentiles confirma que las validaciones capturan los focos principales de generación de viajes, mientras que las salidas reflejan un proceso espacialmente más distribuido, consistente con un patrón de desplazamiento desde pocos orígenes hacia múltiples destinos urbanos.

**Cuadro 1: Estadísticas descriptivas del vector de características  $X_{gt}$**

Variable	Media	Desv. Est.	P25	P50	P75	P95
<b>PIB Millones (COP constantes)</b>	29.059	27.746	2.758	23.008	46.792	82.310
<b>Radiancia VIIRS promedio</b>	49,32	24,68	31,67	52,59	66,33	86,38
<b>Días despejados (VIIRS)</b>	10,87	6,34	6,00	10,00	15,00	22,00
<b>Población residente (Censo 2018)</b>	3,385	3,200	327	2,708	5,459	9,507
<b>Viviendas</b>	1.194	1.091	124	1.042	1.925	3.176
<b>N. Empresas</b>	134,33	204,03	10,00	80,00	188,00	423,00
<b>Salidas TM 00:00–06:00</b>	33,05	35,94	9,06	21,83	44,51	97,44
<b>Salidas TM 06:00–12:00</b>	343,14	309,19	114,20	261,86	487,61	898,37
<b>Salidas TM 12:00–18:00</b>	476,70	503,23	120,49	255,14	698,08	1.584,04
<b>Salidas TM 18:00–24:00</b>	455,46	585,00	65,27	154,98	687,07	1.750,95
<b>Validaciones TM 00:00–06:00</b>	301,33	568,58	13,11	67,92	314,89	1.500,77
<b>Validaciones TM 06:00–12:00</b>	618,43	790,64	104,33	287,12	837,52	2.324,05
<b>Validaciones TM 12:00–18:00</b>	402,42	460,68	99,45	248,77	538,22	1.284,82
<b>Validaciones TM 18:00–24:00</b>	105,43	115,78	26,85	65,71	146,24	340,50

Notas: El PIB observado por celda se encuentra expresado en millones de pesos constantes. Las variables de movilidad (salidas y validaciones de TransMilenio) están medidas en miles de personas por trimestre. Las variables de población y viviendas se reportan en conteos absolutos. La radiancia VIIRS corresponde a valores calibrados del producto VNL VCMCFG. Los días despejados provienen de las banderas operacionales del sensor VIIRS-DNB.

Fuente:

Elaboración: Secretaría Distrital de Desarrollo Económico - Observatorio de Desarrollo Económico de Bogotá.

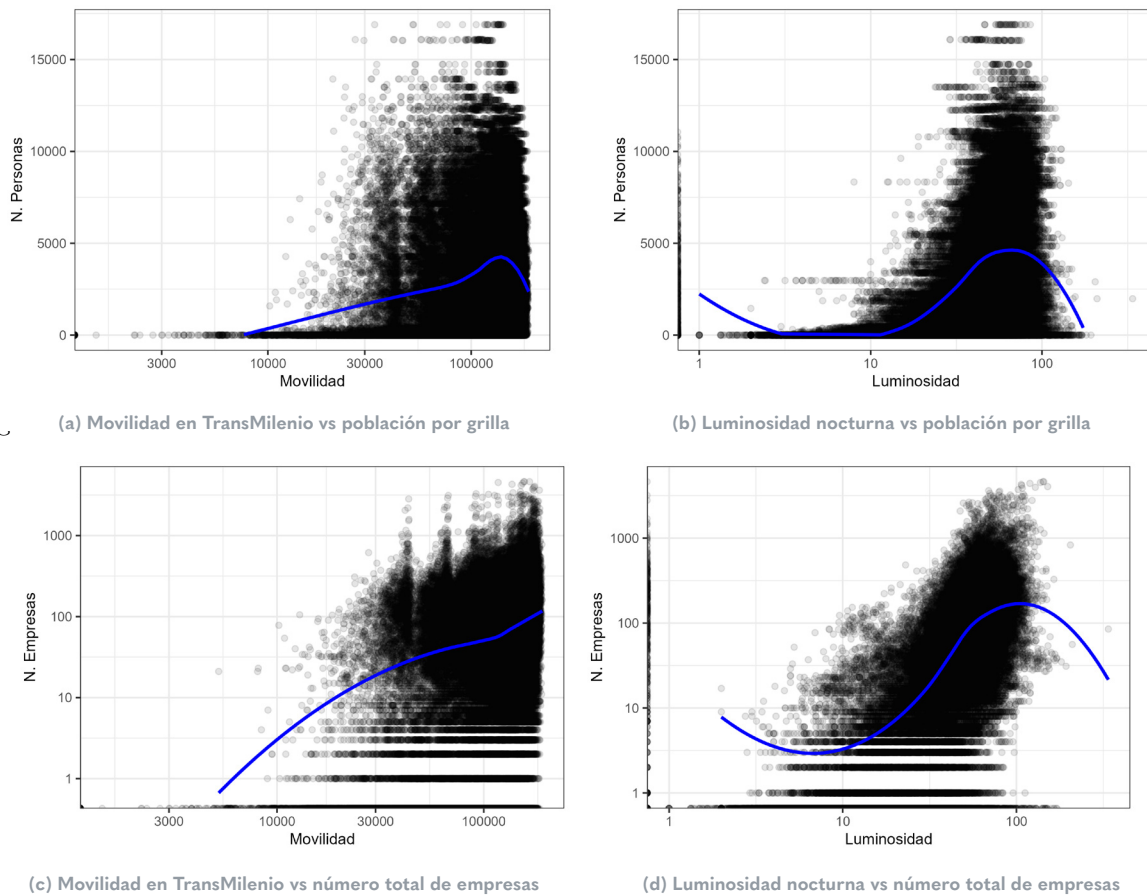
Dado que la población es el insumo que ancla la espacialización del PIB desde la perspectiva de demanda, resulta central examinar cómo se correlaciona con las señales operativas que capturan intensidad urbana —movilidad del sistema de transporte y radiancia nocturna. La representación conjunta de número de personas con movilidad y con luminosidad permite evaluar si las áreas con mayor concentración poblacional exhiben mayor actividad funcional diurna y nocturna, y si estas asociaciones son consistentes con los patrones espaciales previamente observados. En última instancia, estas relaciones bivariadas ilustran las regularidades que el modelo aprovechará para aprender el vínculo entre variaciones en movilidad y luminosidad y el nivel de PIB que se distribuye espacialmente a partir de la estructura demográfica.

La Figura 3 presenta las relaciones bivariadas entre movilidad, luminosidad, población residente y número de empresas a nivel de grilla. El Panel 3a muestra que la dispersión entre población y movilidad es amplia: existen grillas con población baja y niveles elevados de movilidad, lo cual es consistente con zonas de vocación comercial o logística donde la afluencia es alta pese a la baja residencia. También se observan grillas donde altos conteos poblacionales se asocian con altos niveles de movilidad, reflejando áreas residenciales que funcionan como emisoras o receptoras significativas de viajes. En contraste, los puntos con población elevada y movilidad muy baja son escasos, lo que sugiere que, en la mayoría de los casos, los residentes en zonas densas participan en patrones de desplazamiento que se materializan en el sistema de transporte público.

En el Panel 3b, la relación entre luminosidad y población sigue un patrón similar: aunque existe una asociación positiva en el grueso de la distribución, se identifican celdas con alta población y luminosidad moderada y, simultáneamente, celdas con alta luminosidad y baja población, típicamente asociadas a corredores comerciales. En los Paneles 3c y 3d, las asociaciones entre presencia empresarial formal, movilidad y luminosidad muestran gradientes más definidos. En el caso de movilidad y número de empresas, la relación es crecientemente positiva: las grillas con mayor intensidad de transporte concentran sistemáticamente más establecimientos, lo que es coherente con el rol de los nodos de TransMilenio como articuladores de áreas comerciales y zonas de empleo. Existen, sin embargo, celdas con alta movilidad y pocos establecimientos, compatibles con áreas predominantemente residenciales que funcionan como puntos de emisión de viajes.

Finalmente, la asociación entre luminosidad y número de empresas es aún más marcada, con incrementos pronunciados en presencia empresarial a niveles medios de radiancia y una estabilización en los tramos superiores, consistente con el hecho de que la luminosidad captura concentración de infraestructura y actividad económica más que densidad residencial. Esto sugiere que movilidad y luminosidad proporcionan señales complementarias para caracterizar heterogeneidad intraurbana y que ambas exhiben regularidades sistemáticas a partir de las cuales el modelo puede identificar variaciones espaciales en la actividad económica, mientras que su relación con la población está mediada por el uso del suelo y la función económica local.

**Figura 3: Relaciones bivariadas entre luminosidad, movilidad, población y actividad empresarial**



Fuente:

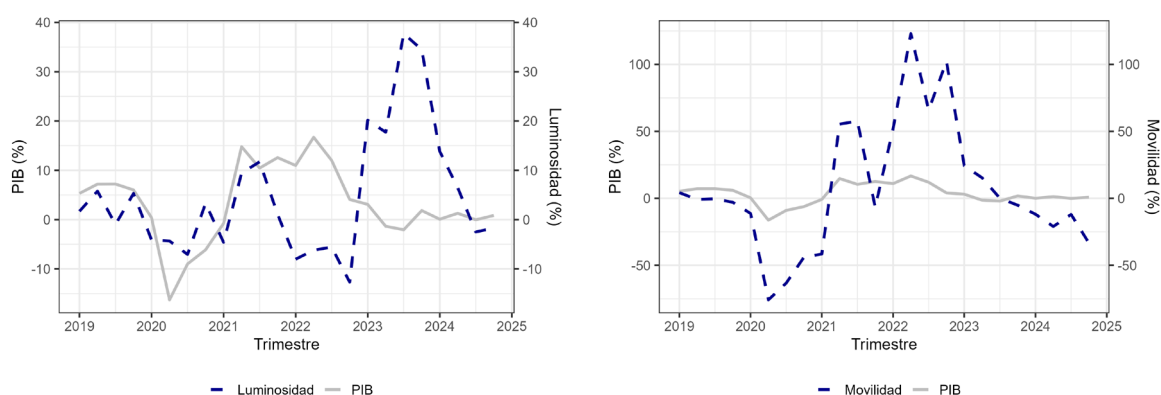
Elaboración: Secretaría Distrital de Desarrollo Económico - Observatorio de Desarrollo Económico de Bogotá.

A continuación, en la Figura 4 se explora en qué medida las variaciones agregadas de luminosidad nocturna y movilidad capturan —con distinta elasticidad— las fluctuaciones de la actividad económica agregada. Aunque la serie de luminosidad exhibe una volatilidad considerablemente mayor que la del PIB, ambas mantienen una trayectoria sincronizada entre el primer semestre de 2019 y el tercer trimestre de 2022. Durante este periodo, los movimientos de la radiancia nocturna replican la dirección del ciclo económico: la contracción abrupta del PIB en los trimestres afectados por la pandemia (2020Q1–2020Q2) coincide con caídas pronunciadas en luminosidad; la recuperación iniciada en 2021Q2 se refleja en aumentos equivalentes en radiancia; y, aunque la luminosidad muestra una caída más marcada en 2022Q1 que el PIB, ambas series retoman un patrón de co-movimiento positivo durante los trimestres siguientes de 2022.

La mayor amplitud de las variaciones en luminosidad confirma que esta señal es más elástica que el PIB ante cambios en el entorno urbano, pero la correspondencia direccional persistente sugiere que, hasta 2022, la radiancia operaba como un indicador consistente del ciclo económico urbano. Esta relación cambia a partir del tercer trimestre de 2022. Desde ese punto, la luminosidad se desacopla parcialmente del PIB y registra oscilaciones abruptas que no encuentran contrapartida en el agregado macroeconómico: repuntes atípicos hacia finales de 2022 y, especialmente, en 2023Q1–Q4, superan ampliamente la magnitud de las variaciones del PIB, indicando que la radiancia incorpora dinámicas operativas o infraestructurales que no se reflejan en el producto agregado.

Por otra parte, la movilidad exhibe un patrón análogo pero con rupturas más tempranas. Entre 2019 y 2021, movilidad y PIB se mueven en la misma dirección, aunque con amplitudes muy superiores en movilidad. La contracción de más del 50 % en 2020Q2 supera ampliamente la caída del PIB, seguida de una recuperación acelerada durante 2021. Sin embargo, a partir de 2022, la movilidad se vuelve altamente volátil y pierde alineación con el ciclo económico, reflejando cambios modales post-pandemia, ajustes operativos y variabilidad intraurbana no capturada por el PIB. En conjunto, estas rupturas estructurales posteriores a 2022 pueden representar un desafío en el modelado de la relación PIB–luminosidad y PIB–movilidad a lo largo del tiempo, aprovechando la fase de alta sincronía previa y ajustando el peso relativo de estas señales cuando su elasticidad deja de ser proporcional al agregado macroeconómico.

**Figura 4: Crecimiento del PIB comparado con luminosidad nocturna y movilidad**



(a) Tasa de crecimiento del PIB y de la luminosidad nocturna

(b) Tasa de crecimiento del PIB y de la movilidad en TransMilenio

Fuente:

Elaboración: Secretaría Distrital de Desarrollo Económico - Observatorio de Desarrollo Económico de Bogotá.

### 3. METODOLOGÍA

---

En esta sección se describe la metodología empleada para estimar el PIB a nivel de grilla para Bogotá utilizando las fuentes de luminosidad nocturna, movilidad y estructura urbana previamente presentadas. El primer reto metodológico consiste en definir un mecanismo consistente para espacializar el PIB oficial, cuya unidad de observación es la ciudad, hacia una grilla regular de 450×450 m. La literatura ha abordado este problema mediante asignaciones uniformes del PIB entre celdas o mediante reglas proporcionales basadas en luminosidad nocturna. Sin embargo, estos enfoques no resultan adecuados para un esquema de estimación orientado a la demanda, en el cual el nivel del PIB debe reflejar la distribución espacial de los hogares y su participación agregada en el gasto.

En este estudio, la espacialización del PIB se realiza utilizando la población residente como variable de reparto, aprovechando que existe información censal de alta resolución para Bogotá. Bajo este enfoque, se asume que cada residente contribuye, en promedio, con el mismo nivel de PIB per cápita dentro de la ciudad, independientemente de la heterogeneidad socioeconómica o sectorial entre zonas urbanas. Este supuesto —estándar en ejercicios de desagregación subnacional— permite anclar el nivel del PIB en la estructura demográfica observada y es coherente con un enfoque de demanda en el que los hogares constituyen el punto de partida del gasto agregado.

#### (i) Prior y *pseudo-target*

Sea  $g$  el índice de celdas de la grilla y  $t$  los trimestres. Para cada ciudad  $c$  se define un ancla de nivel consistente con las cuentas nacionales mediante el PIB per cápita oficial:

$$\text{GDPpc}_{ct} = \frac{\text{GDP}_{ct}}{\text{Pop}_c}.$$



Este nivel se distribuye espacialmente utilizando la población del Censo 2018, tratada como invariante en todo el horizonte analizado debido a la ausencia de información subanual de alta resolución. La asignación inicial del PIB por celda se obtiene multiplicando el nivel per cápita por la población residente en cada unidad espacial:

$$\widetilde{Y}_{gt} = \text{GDPpc}_{ct} \times \text{Pop}_g^{(2018)}.$$

La construcción de  $\widetilde{Y}_{gt}$  cumple dos funciones. Primero, garantiza consistencia contable entre los agregados oficiales y la desagregación espacial del PIB. Segundo, establece una línea base sobre la cual las señales de alta frecuencia —luminosidad, movilidad y variables estructurales— capturan la heterogeneidad intraurbana y la variación temporal. En consecuencia, el modelo no estima niveles absolutos desde cero, sino que aprende desviaciones espacio-temporales alrededor del *pseudo-target* anclado en población.

## (ii) Modelo, features y entrenamiento

La estimación del PIB por grilla se formula mediante un modelo de red neuronal *feedforward* que aproxima la función no lineal entre el vector de características  $X_{gt}$  y el *pseudo-target*  $\widetilde{Y}_{gt}$ :

$$f : X_{gt} \longrightarrow \widehat{Y}_{gt}.$$

El vector  $X_{gt}$  consolida señales exógenas de alta frecuencia y atributos estructurales invariables que reflejan heterogeneidad urbana. Las variables incluidas corresponden estrictamente a las disponibles en la Tabla 1 y se agrupan en cuatro bloques:

1. **Radiancia VIIRS-DNB:** nivel promedio trimestral de radiancia y número de días despejados, este último utilizado como indicador de confiabilidad radiométrica.
2. **Movilidad en TransMilenio:** salidas y validaciones trimestrales en cuatro franjas horarias (00:00–06:00, 06:00–12:00, 12:00–18:00 y 18:00–24:00), expresadas en miles de personas. Estas métricas capturan intensidad funcional diurna y nocturna y permiten incorporar el componente horario de la dinámica urbana.
3. **Estructura urbana:** población residente (Censo 2018) y número de viviendas por celda, ambas tratadas como invariantes en el período analizado.
4. **Controles espaciales:** coordenadas X e Y del centroide de cada celda, utilizadas para modelar la dependencia espacial implícita en la estructura urbana. Estos controles permiten absorber gradientes espaciales de baja frecuencia y capturar la proximidad relativa entre celdas vecinas sin imponer una estructura paramétrica específica.

El modelo se entrena bajo un esquema de optimización diseñado para explorar de manera exhaustiva la sensibilidad del estimador frente a distintas configuraciones arquitectónicas y combinaciones de señales. En el caso de las redes neuronales, se evalúan cinco especificaciones alternativas del vector de características, cada una asociada a un barrido sistemático de hiperparámetros. El espacio de hiperparámetros para las redes neuronales incluye variación en el número de capas ocultas, número de neuronas por capa, tasa de aprendizaje, tasa de *dropout*, tamaño del *batch* y funciones de activación. Además, se utiliza un criterio de detención anticipada basado en desempeño de validación: si la pérdida ponderada no mejora durante diez iteraciones consecutivas, el entrenamiento se interrumpe.

### (iii) Validación y selección de modelo

El entrenamiento y la validación se realizan bajo un esquema temporal estricto. La base trimestral 2012Q2–2023Q4 se divide en una partición 70–30 % en orden cronológico, evitando fugas de información. Toda la calibración interna —incluida la selección de hiperparámetros mediante validación *rolling origin*— se ejecuta exclusivamente dentro de dicho intervalo. Para cada modelo se calculan predicciones agregadas a nivel de ciudad:

$$\widehat{Y}_{ct} = \sum_g \widehat{Y}_{gt}.$$

Las predicciones agregadas  $\widehat{Y}_{ct}$  se contrastan con las proyecciones oficiales del PIB de 2024, lo que permite evaluar coherencia tanto en niveles como en tasas de crecimiento. Los modelos que superan esta evaluación generan posteriormente la senda trimestral del PIB por grilla para 2025, construida únicamente a partir de señales exógenas de alta frecuencia y de la estructura aprendida en el período histórico.

Los modelos se clasifican inicialmente según el RMSE en la base de prueba, reteniéndose aquellos con un error inferior al 0.8 % de la media del PIB a nivel de grilla. Entre este subconjunto, se seleccionan los diez modelos que minimizan la diferencia absoluta entre la tasa de crecimiento oficial del PIB de la ciudad y la tasa de crecimiento implícita en  $\widehat{Y}_{ct}$  durante 2024. Con estos diez modelos se construye un ensemble descriptivo del crecimiento agregado, utilizado como rango de dispersión y no como intervalo de confianza.

## 4. RESULTADOS

---

En esta sección se presentan los resultados del ejercicio de predicción del PIB a nivel de grilla. El entrenamiento consideró un conjunto amplio de arquitecturas y especificaciones, diseñado para evaluar la sensibilidad del modelo frente a distintas combinaciones de señales, hiperparámetros y estructuras funcionales. En total se estimaron siete familias de modelos: redes neuronales, gradiente boosting (GBM) y bosques aleatorios (RF). Dentro de las redes neuronales se exploraron cinco combinaciones alternativas de variables explicativas, cada una asociada a un barrido sistemático de hiperparámetros. Esta estrategia permitió evaluar la robustez del mapeo entre las señales exógenas de alta frecuencia y el PIB por grilla, y facilitó la selección de modelos consistentes en predicción agregada y en tasas de crecimiento.

El espacio de hiperparámetros para las redes neuronales incluyó variación en: número de capas ocultas, número de neuronas por capa, tasa de aprendizaje, tasa de *dropout*, tamaño del batch (16 o 32) y funciones de activación. Para cada combinación de variables y conjunto de hiperparámetros se entrenaron aproximadamente 16.000 modelos, alcanzando un total cercano a 16.000 entrenamientos efectivos. Cada modelo se calibró mediante un máximo de 350–450 iteraciones (*epochs*), con un criterio de detención anticipada basado en el comportamiento del error de validación: si después de diez iteraciones consecutivas no se observaba mejora en la pérdida ponderada, el entrenamiento se detenía.

Adicionalmente, el proceso de entrenamiento se estructuró en un esquema temporal estricto. La base de datos se construyó a frecuencia trimestral desde 2012Q2 hasta 2023Q4, período utilizado para estimar y calibrar los modelos. La partición entrenamiento–prueba siguió la regla tradicional 70–30 %, manteniendo la integridad temporal de la serie y evitando fugas de información. Toda la calibración interna del modelo, incluyendo la selección de hiperparámetros mediante validación *rolling origin*, se realizó exclusivamente dentro del intervalo 2012–2023. Para evaluar desempeño fuera de muestra, se implementó una validación externa utilizando las proyecciones oficiales del PIB agregadas para 2024, que permiten contrastar el crecimiento implícito en las predicciones agregadas  $\widehat{Y}_{ct}$  frente al crecimiento observado. Finalmente, cada modelo ganador se utilizó para producir predicciones trimestrales del PIB a nivel de grilla durante 2025, generando así una senda prospectiva basada únicamente en señales de alta frecuencia y en la estructura aprendida durante el período histórico.

El análisis posterior se estructuró en dos etapas. Primero, se documenta el espacio de búsqueda utilizado y se reportan los parámetros sobre los cuales se generó variación sistemática. Segundo, se identifican y describen los modelos con mejor desempeño según dos criterios: (i) el RMSE de las predicciones agregadas a nivel ciudad, y (ii) la proximidad entre la tasa de crecimiento agregada implícita en las predicciones y la tasa de crecimiento oficial. La pre-presentación de resultados se concentra en los cinco modelos superiores en cada criterio y en el modelo final seleccionado, incluyendo métricas de ajuste, estabilidad temporal y predicciones agregadas por ciudad y por trimestre.

## 4.1. ESPACIO DE BÚSQUEDA

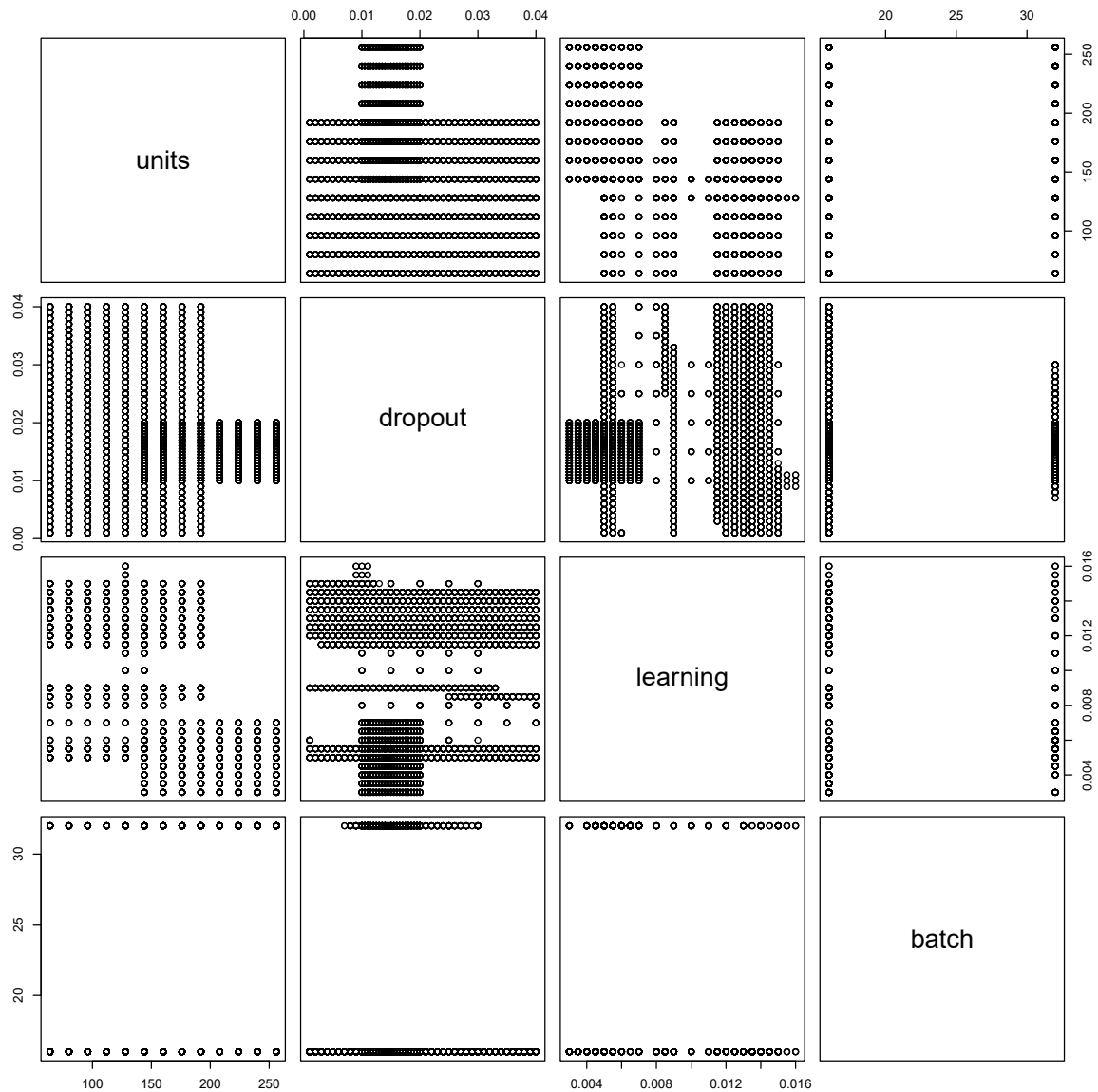
El espacio de búsqueda de hiperparámetros incluyó cuatro dimensiones principales: número de unidades por capa (*units*), tasa de abandono (*dropout*), tasa de aprendizaje (*learning rate*) y tamaño del lote (*batch*). La primera ola de entrenamiento se planteó como un barrido amplio sobre todo el soporte paramétrico: las arquitecturas variaron entre 64 y 196 neuronas por capa, la tasa de abandono entre 0.01 y 0.04, la tasa de aprendizaje entre 0.004 y 0.016, y el tamaño del lote entre 16 y 32. Este ejercicio inicial permitió caracterizar de manera exploratoria el espacio de modelos factibles y estimar sensibilidad del error predictivo a variaciones amplias en cada dimensión.

A partir de los resultados de esta primera ola se identificaron zonas del espacio paramétrico donde el RMSE tendía a ser menor. Los modelos con mejor desempeño se concentraban en configuraciones con tasas de abandono entre 0.01 y 0.02 y tasas de aprendizaje inferiores a 0.006. Dentro de esas combinaciones, las arquitecturas con 128 neuronas o más por capa mostraban, en promedio, mayor estabilidad en validación. Sobre esa base se diseñó una segunda ola de entrenamiento, enfocada deliberadamente en los rangos que habían mostrado mejor relación señal–error. El barrido mantuvo *dropout* entre 0.01 y 0.02, restringió la tasa de aprendizaje al intervalo 0.003–0.006 y expandió el número de neuronas hacia la franja 128–256. El tamaño del lote se mantuvo en 16 y 32 tras verificar que no existían diferencias sistemáticas en desempeño entre ambas opciones.

La Figura 5 refleja esta secuencia de entrenamiento. La mayor densidad de puntos en los rangos de *dropout* 0.01–0.02 y de *learning rate* por debajo de 0.006 corresponde a la segunda ola del barrido, enfocada en las combinaciones donde se habían observado los menores errores en la fase inicial. La dispersión más amplia en el número de unidades —especialmente entre 150 y 256— proviene de la ampliación deliberada del espacio arquitectural en la etapa refinada del entrenamiento. Por el contrario, la ausencia de modelos con *dropout* superiores a 0.02 o tasas de aprendizaje por encima de 0.01 refleja que esas combinaciones no fueron retomadas después del barrido inicial debido a su menor desempeño relativo.

Los paneles correspondientes al tamaño del lote muestran dos conglomerados definidos (16 y 32), sin patrón adicional, lo que indica que ambas opciones se mantuvieron activas a lo largo de ambas etapas sin inducir estructura visible en el resto del espacio de búsqueda. Finalmente, los modelos seleccionados como candidatos finales se ubican prácticamente por completo dentro del subconjunto de hiperparámetros afinado durante la segunda ola: *dropout* entre 0.01 y 0.02, tasa de aprendizaje entre 0.003 y 0.006 y arquitecturas con 128 neuronas o más por capa.

**Figura 5: Exploración del espacio de hiperparámetros del modelo**



Fuente:

Elaboración: Secretaría Distrital de Desarrollo Económico - Observatorio de Desarrollo Económico de Bogotá.



## 4.2. DESEMPEÑO DEL MODELO

El Cuadro 1 presenta el desempeño de los diez modelos con menor error dentro de todo el conjunto de configuraciones entrenadas. El objetivo no es comparar estos modelos entre sí, sino mostrar que, dentro del espacio total de búsqueda, existe un conjunto amplio de especificaciones capaces de reproducir el PIB con alta precisión. El RMSE —expresado como proporción del PIB— cuantifica la desviación promedio entre la predicción por grilla y el *pseudo-target* asignado en cada trimestre. columna ‘Diferencia de Tasas’ evalúa un criterio distinto: la discrepancia absoluta entre la tasa de crecimiento oficial del PIB y la tasa implícita en las predicciones agregadas para 2024. Dado que el interés principal es la coherencia del crecimiento agregado, esta métrica resume qué tan cerca se encuentran las predicciones del ciclo económico observado. Las diferencias se mantienen por debajo de 1 punto porcentual en todos los casos, lo que implica que, aun con múltiples configuraciones, los modelos producen trayectorias agregadas del PIB consistentes con las cifras oficiales.

**Cuadro 2: Desempeño de los modelos predictivos**

Modelo	RMSE ( % )	Diferencia de Tasas ( % )
Modelo 1	0,624	0,658
Modelo 2	0,627	0,679
Modelo 3	0,668	0,726
Modelo 4	0,734	0,982
Modelo 5	0,762	0,920
Modelo 6	0,767	0,926
Modelo 7	0,788	0,367
Modelo 8	0,789	0,569
Modelo 9	0,796	0,665
Modelo 10	0,796	0,370

**Nota:** RMSE corresponde al error cuadrático medio expresado como proporción de la variable de resultado. La columna ‘Diferencia de Tasas’ muestra la variación promedio entre la tasa de crecimiento del PIB y la tasa predicha para el año 2024.

Fuente:

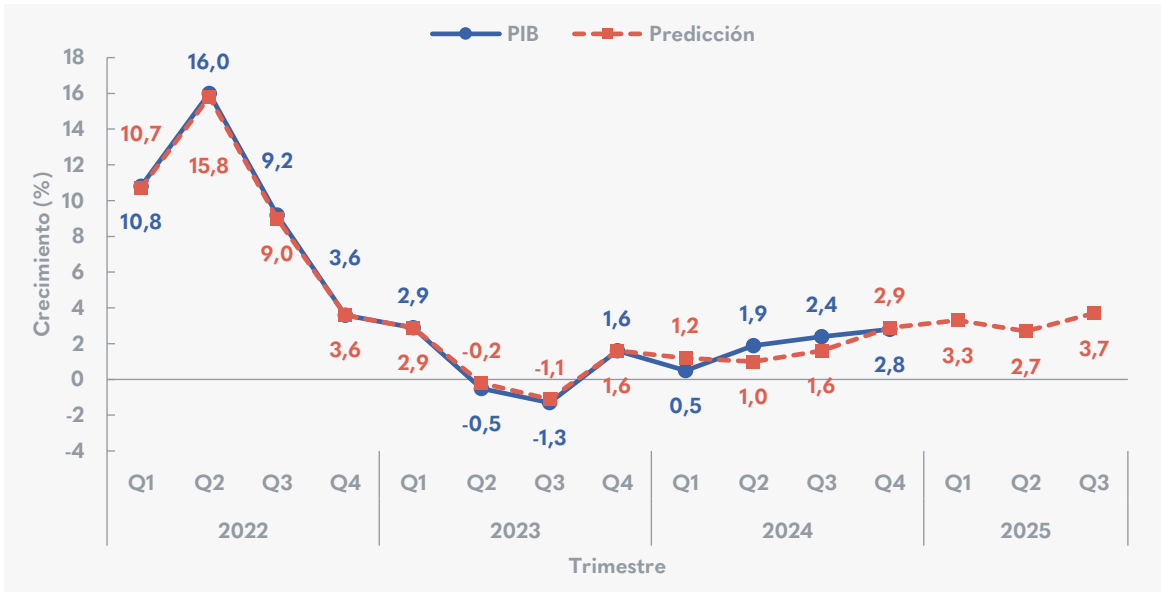
Elaboración: Secretaría Distrital de Desarrollo Económico - Observatorio de Desarrollo Económico de Bogotá.

La Figura 6 presenta la comparación entre la tasa de crecimiento oficial del PIB y la tasa implícita en las predicciones agregadas para el período 2022Q1–2025Q3. La línea punteada roja corresponde al promedio de los diez modelos con mejor desempeño, mientras que la banda gris representa el rango completo de predicciones (mínimo–máximo) dentro de este conjunto, funcionando como un abanico de variación plausible más que como un intervalo de confianza estadístico. Entre 2022 y 2023, la franja gris es estrecha, lo que refleja la baja dispersión entre modelos: este comportamiento es coherente con los RMSE inferiores al 1 % observados en la etapa de validación y evidencia que, en esta ventana, los modelos producen predicciones muy similares entre sí.

En 2024 la banda gris se amplía levemente, pues el modelo ya no cuenta con observaciones efectivas del PIB para este año y debe inferir la dinámica exclusivamente a partir de señales de luminosidad y movilidad. Aun así, las predicciones —tanto el promedio como el rango— se alinean estrechamente con las cifras oficiales disponibles, tal como se documenta en la columna de “Diferencia de Tasas” del Cuadro 1. A partir de 2025 todas las series corresponden a predicciones puras, dado que aún no existen cifras oficiales del PIB para contrastar. En esta sección la banda gris se ensancha de manera notable, reflejando mayor varianza entre modelos: la amplitud responde a dos factores. Primero, a la ausencia de un ancla oficial que discipline la trayectoria trimestral. Segundo, al hecho de que las señales de luminosidad y movilidad muestran fluctuaciones más pronunciadas en los trimestres recientes, lo que induce mayor dispersión en las proyecciones.

A pesar de ello, el patrón agregado conserva elementos estructurales relevantes: se mantiene la estacionalidad típica de repuntes en el segundo y tercer trimestre, y el crecimiento promedio proyectado para 2025Q2–Q3 (3.3 % y 2.7 %) es coherente con los pronósticos macroeconómicos externos. Para 2025Q3, por ejemplo, el promedio del modelo (3.7 %) es prácticamente idéntico a la estimación preliminar publicada por el DANE (3.6 % para Colombia), y se ubica dentro del rango de pronósticos recientes del Indicador de Actividad Económica del Banco de la República. El abanico de predicciones para ese trimestre oscila entre 2 % y 5 %, lo que sugiere que, incluso con mayor incertidumbre hacia adelante, el modelo genera un conjunto de trayectorias razonables y consistentes con la información macroeconómica disponible.

**Figura 6: Crecimiento económico estimado vs. PIB observado**



Nota: La figura presenta el promedio de las predicciones generadas por los cinco modelos con mejor desempeño, comparadas con la tasa de crecimiento observada del PIB. La línea azul corresponde al crecimiento observado; la línea roja discontinua, al crecimiento estimado; y la banda gris, a la desviación estándar entre las predicciones de los modelos seleccionados.

Fuente:

Elaboración: Secretaría Distrital de Desarrollo Económico - Observatorio de Desarrollo Económico de Bogotá.

## 5. CONCLUSIONES

---

El estudio muestra que es posible generar estimaciones trimestrales del PIB a escala intraurbana empleando exclusivamente señales exógenas de alta frecuencia—radiancia nocturna, movilidad del sistema de transporte masivo y estructura demográfica—sin recurrir a registros administrativos en la etapa predictiva. La combinación entre un *pseudo-target* anclado en población y un modelo no lineal entrenado con validación temporal estricta produce estimaciones espacialmente coherentes y temporalmente estables. Los resultados indican que, aun en presencia de heterogeneidad urbana y variaciones operativas en las señales satelitales y de movilidad, el modelo reproduce con alta fidelidad la trayectoria agregada del PIB de Bogotá durante el período observado.

La evaluación fuera de muestra confirma la capacidad predictiva del enfoque. Las diferencias entre la tasa de crecimiento oficial del PIB y las tasas implícitas en las predicciones agregadas se mantienen por debajo de un punto porcentual en 2024, lo que sugiere una alineación estrecha entre las señales de alta frecuencia y el ciclo económico urbano. Además, el modelo preserva patrones estructurales relevantes—como la estacionalidad del crecimiento en el segundo y tercer trimestre—aun en los ejercicios prospectivos para 2025, año para el cual no existe información oficial. Aunque la dispersión entre modelos aumenta en los periodos sin anclas externas, las proyecciones permanecen dentro de rangos compatibles con pronósticos macroeconómicos independientes, lo que refuerza la solidez del procedimiento.

Finalmente, los resultados abren una agenda clara para el desarrollo futuro. Si las variaciones agregadas de movilidad y luminosidad capturan de manera consistente la dinámica del PIB de la ciudad, entonces estas señales pueden emplearse para analizar la evolución económica a escalas territoriales más finas—localidades, barrios o incluso manzanas—y para monitorear cambios en tiempo casi real. La combinación de estas dos fuentes ofrece un insumo diagnóstico valioso para identificar zonas en expansión, áreas en contracción y patrones espaciales de actividad económica que no son visibles en las cuentas oficiales. El siguiente paso natural de este trabajo consiste en descender de la agregación ciudad a una desagregación plenamente espacial, aprovechando la granularidad del panel de grillas para estudiar dinámicas intraurbanas persistentes y emergentes. Este proceso permitirá evaluar heterogeneidad territorial, detectar focos de cambio económico y avanzar hacia sistemas operativos de monitoreo económico urbano basados en señales de alta frecuencia.

## 6. BIBLIOGRAFÍA



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

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. URL: <https://doi.org/10.1073/pnas.1017031108>

Doll, C. N. H., Muller, J.-P., & Morley, J. G. (2006). Night-time imagery as a tool for global mapping of socioeconomic parameters. *Ambio: A Journal of the Human Environment*, 29(3), 157–162. URL: [https://doi.org/10.1639/0044-7447\(2000\)029\[0157:NTIAAT\]2.0.CO;2](https://doi.org/10.1639/0044-7447(2000)029[0157:NTIAAT]2.0.CO;2)

Ehsan, Z.-A., Basher, S. A., Rashid, S., & Uddin, M. R. (2020). Estimation of District-Level GDP Using Night Lights: The Case of Bangladesh (inf. téc.). URL (preprint SSRN): <https://doi.org/10.2139/ssrn.4373887>

Henderson, J. V., Storeygard, A., & Weil, D. N. (2012). Measuring Economic Growth from Outer Space. *American Economic Review*, 102(2), 994–1028. URL: <https://doi.org/10.1257/aer.102.2.994>

Lessmann, C., & Seidel, A. (2016). Regional Inequality, Convergence, and its Determinants—A View from Outer Space. *European Economic Review*, 92, 110–132. URL: <https://doi.org/10.1016/j.eurocorev.2016.11.009>

Liu, H., He, X., Bai, Y., Liu, X., Wu, Y., Zhao, Y., & Yang, H. (2021). Nightlight as a Proxy of Economic Indicators: Fine-Grained GDP Inference around Chinese Mainland via Attention-Augmented CNN from Daytime Satellite Imagery. *Remote Sensing*, 13(11), 2067. URL: <https://doi.org/10.3390/rs13112067>

McCord, G. C., & Rodriguez-Heredia, M. (2022). Nightlights and Subnational Economic Activity: Estimating Departmental GDP in Paraguay. *Remote Sensing*, 14(5), 1150. URL: <https://doi.org/10.3390/rs14051150>

Suleiman, H. (2021). Illuminating the Nile: Estimating Subnational GDP in Egypt using Nighttime Lights and Machine Learning (inf. téc.). URL: <https://doi.org/10.1007/s10708-024-11106-6>

## Abreviaturas

<b>PIB</b>	Producto Interno Bruto.
<b>GDP</b>	Gross Domestic Product.
<b>RMSE</b>	Raíz del Error Cuadrático Medio (Root Mean Squared Error).
<b>GBM</b>	Gradient Boosting Machine.
<b>RF</b>	Random Forest.
<b>Q1, Q2, Q3, Q4</b>	Primer, segundo, tercer y cuarto trimestre.



# Siglas

<b>CIENFI</b>	Centro de Investigación en Economía y Finanzas.
<b>SDDE</b>	Secretaría de Desarrollo Económico de Bogotá.
<b>DANE</b>	Departamento Administrativo Nacional de Estadística.
<b>RUES</b>	Registro Único Empresarial y Social.
<b>TM</b>	TransMilenio.
<b>VIIRS-DNB</b>	Visible Infrared Imaging Radiometer Suite – Day/ Night Band.
<b>VNL VCMCFG</b>	Producto mensual de luces nocturnas VIIRS libre de nubosidad del Earth Observation Group.
<b>EOG</b>	Earth Observation Group.
<b>NOAA</b>	National Oceanic and Atmospheric Administration.
<b>Suomi-NPP</b>	Suomi National Polar-orbiting Partnership.

## Índice de cuadros

<b>Cuadro 1:</b>	Estadísticas descriptivas del vector de características $X_{gt}$	<b>17</b>
<b>Cuadro 2:</b>	Desempeño de los modelos predictivos	<b>27</b>

## Índice de figuras

<b>Figura 1:</b>	Distribución espacial de luminosidad nocturna y población en Bogotá	<b>14</b>
<b>Figura 2:</b>	Validaciones y salidas por troncales de TransMilenio	<b>15</b>
<b>Figura 3:</b>	Relaciones bivariadas entre luminosidad, movilidad, población y actividad empre-sarial	<b>19</b>
<b>Figura 4:</b>	Crecimiento del PIB comparado con luminosidad nocturna y movilidad	<b>20</b>
<b>Figura 5:</b>	Exploración del espacio de hiperparámetros del modelo	<b>26</b>
<b>Figura 6:</b>	Crecimiento económico estimado vs. PIB observado	<b>28</b>



