

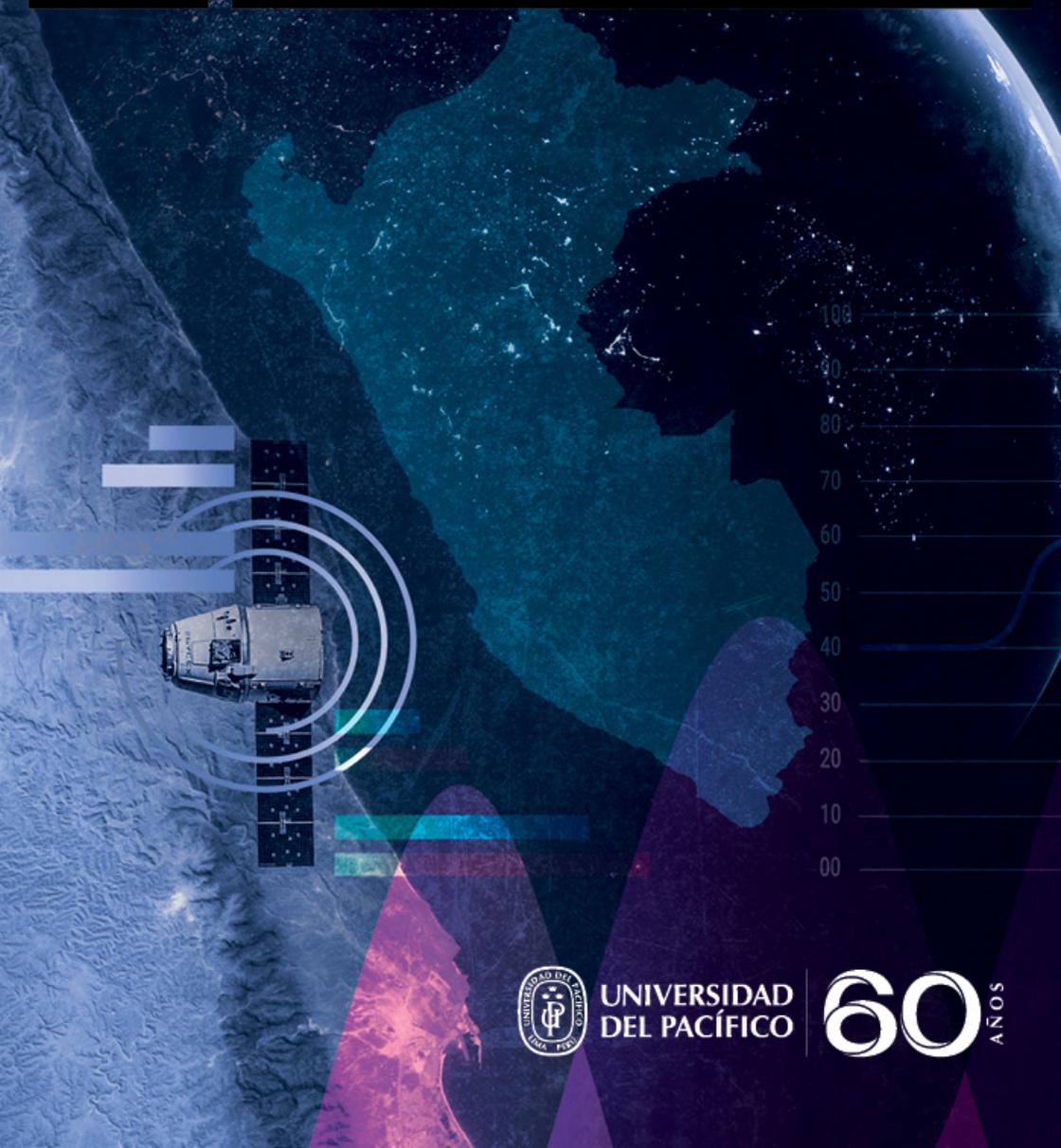
Bruno Seminario y Luis Palomino

20

DOCUMENTO
DE INVESTIGACIÓN

Estimación del PIB a nivel subnacional utilizando datos satelitales de luminosidad:

Perú, 1993-2018



UNIVERSIDAD
DEL PACÍFICO

60
AÑOS

Bruno Seminario y Luis Palomino

20

DOCUMENTO
DE INVESTIGACIÓN

Estimación del PIB a nivel
subnacional utilizando datos
satelitales de luminosidad:
Perú, 1993–2018



UNIVERSIDAD
DEL PACÍFICO

60
AÑOS

© Bruno Seminario y Luis Palomino, 2022

De esta edición:

© Universidad del Pacífico
Jr. Gral. Luis Sánchez Cerro 2141
Lima 15072, Perú

**Estimación del PIB a nivel subnacional utilizando datos satelitales de luminosidad:
Perú, 1993-2018**

Bruno Seminario y Luis Palomino

1.ª edición: mayo de 2022

Diseño de la carátula: Ícono Comunicadores

ISBN ebook: 978-9972-57-493-1

Hecho el Depósito Legal en la Biblioteca Nacional del Perú: 2022-04387

doi: <http://dx.doi.org/10.21678/978-9972-57-493-1>

Disponible en fondoeditorial.up.edu.pe

BUP

Seminario, Bruno

Estimación del PIB a nivel subnacional utilizando datos satelitales de luminosidad:
Perú, 1993-2018 / Bruno Seminario y Luis Palomino. -- 1a edición. -- Lima: Universidad
del Pacífico, 2022.

61 p. -- (Documento de investigación ; 20)

1. Producto interno bruto--Perú--Procesamiento de datos

2. Desarrollo regional--Perú--Procesamiento de datos

3. Desigualdad económica--Perú--Procesamiento de datos

I. Palomino, Luis

II. Universidad del Pacífico (Lima)

339.31 (SCDD)

La Universidad del Pacífico no se solidariza necesariamente con el contenido de los trabajos que publica. Prohibida la reproducción total o parcial de este texto por cualquier medio sin permiso de la Universidad del Pacífico.

Derechos reservados conforme a ley.

Agradecimientos

Agradecemos los valiosos aportes de Luan Sánchez y Ronald Cueva. También agradecemos los valiosos y enriquecedores comentarios de María Alejandra Zegarra, Alejandro Sánchez, Juan Mendoza, María Amparo Cruz-Saco, Marc Badia Miro y Efraín Gonzales de Olarte; aquellos realizados por los dos revisores anónimos y, finalmente, agradecemos también a los participantes del VII Congreso Latinoamericano de Historia Económica 2022 por sus aportes.

Índice

1. Introducción	9
2. Datos y metodología	13
2.1 Datos	13
2.2 Metodología	18
3. Estimación del PIB a nivel subnacional	25
3.1 Estimación del PIB a nivel de distritos	25
3.2 Estimación del PIB a nivel de provincias	32
4. Tendencias recientes de desarrollo regional	37
4.1 Distribución espacial del PIB pc y PIB por km ²	37
4.2 Desigualdad regional	41
4.3 Desarrollo regional	45
5. Conclusiones	49
6. Limitaciones y futuras líneas de investigación	53
Referencias	57
Anexos	61

1. Introducción

El producto interno bruto (PIB) es la variable más importante en los análisis de crecimiento económico. Sin embargo, a menudo se cometen errores de medición, en especial en los países en vías de desarrollo. Estos errores se cometen porque, en comparación con los países desarrollados, en muchos países en desarrollo la presencia del sector informal es mayor, el grado de integración económica y la igualdad de precios entre regiones son menores, y la infraestructura estadística del gobierno es más débil (Henderson, Storeygard, & Weil, 2012).

De otro lado, los estudios regionales sugieren que el desempeño económico en América Latina es heterogéneo. De acuerdo con la Cepal (2015), los elevados niveles de riqueza y crecimiento económico de la última década se concentran en pocas regiones; en particular, se encuentran en las regiones con actividades relacionadas con los servicios y la minería. Por ello, para identificar las regiones de un país que requieren incrementar sus niveles de riqueza y crecimiento económico, es necesario construir estadísticas de actividad económica a nivel de subdivisiones territoriales. Sin embargo, la recopilación de información socioeconómica, en especial los niveles de ingresos, resulta costosa y también problemática, debido a sus deficiencias en el proceso de recopilación de datos.

En este contexto, la conexión entre los niveles de luminosidad (medida por las luces de noche observadas desde el espacio) y las áreas urbanas ha motivado el uso de las imágenes satelitales de luminosidad como un indicador de actividad económica (Levin *et al.*, 2020; Ghosh *et al.*, 2010). Diferentes estudios han encontrado que los niveles de luminosidad están correlacionados con la actividad económica (Amavilah, 2018; Levin & Zhang, 2017; Pinkovskiy & Sala-i-Martin, 2016; Nordhaus & Chen, 2015; Michalopoulos & Papaioannou 2013; Henderson *et al.*, 2012; Ghosh *et al.*, 2010). Sin embargo, también se

han encontrado grandes diferencias en los niveles de luminosidad en países con niveles similares de PIB (Levin & Zhang, 2017; Kyba *et al.*, 2017; Henderson *et al.*, 2012).

En una investigación reciente, Levin *et al.* (2020) indican que el nivel de luminosidad permite tres tipos de análisis económico: (1) estimar el PIB a diferentes escalas espaciales; (2) estimar las tasas de crecimiento del PIB con alta frecuencia temporal (Bennie, Davies, Duffy *et al.*, 2014; Bennie, Davies, Inger, & Gastón, 2014); (3) estimar el PIB en regiones que tienen informes deficientes o no tienen informes sobre estadísticas económicas (Henderson *et al.*, 2012).

Por un lado, el nivel de luminosidad se puede utilizar para analizar el PIB a diferentes escalas espaciales. Este análisis consiste en desagregar los datos del PIB nacional en cuadrículas espaciales. Este ejercicio permitió generar por primera vez un mapa del PIB con una resolución de 5 km para 11 países de la Unión Europea y los Estados Unidos (Doll, Muller, & Morley, 2006). Posteriormente, con el apoyo de datos auxiliares y un mapa de densidad de población, se generó un mapa del PIB global con una resolución de 1 km, mostrando que Singapur tenía la densidad de PIB más alta (Ghosh *et al.*, 2010). El nivel de luminosidad también se puede utilizar como *proxy* del PIB para estimar la riqueza, lo que permite mapear fenómenos económicos regionales como la desigualdad (Elvidge *et al.*, 2012; Xu *et al.*, 2015) y la pobreza (Elvidge *et al.*, 2009; Wang, Cheng, & Zhang, 2012; Yu *et al.*, 2015; Jean *et al.*, 2016). Henderson *et al.* (2018) muestran que la geografía física (como el clima, biomas, topografía, etc.) tiene una fuerte influencia en la distribución espacial de la actividad económica. Sin embargo, existen diferencias en la importancia relativa de las variables agrícolas y comerciales para explicar la variabilidad espacial del nivel de luminosidad entre los países desarrollados y los países en vías de desarrollo.

Por otro lado, el nivel de luminosidad también se puede utilizar para corregir las estadísticas de países en vías de desarrollo. Este ejercicio está basado en modelos econométricos que consideran el PIB real (o tasa de crecimiento del PIB) como una combinación lineal de PIB estadístico (o tasa de crecimiento del PIB) y el PIB estimado (o tasa de crecimiento del PIB) derivado de imágenes satelitales de luminosidad (Chen & Nordhaus, 2011; Henderson *et al.*, 2011, 2012). Este análisis ha permitido concluir que la tasa de crecimiento del PIB real de China es más alta que los valores de las estadísticas oficiales (Clark, Pinkovskiy, & Sala-i-Martin, 2017).

En América Latina, a pesar de que los países de la región no cuentan con datos de calidad de PIB a nivel de subdivisiones administrativas territoriales, los estudios que han utilizado las imágenes satelitales de luminosidad para medir la actividad económica son escasos (Guerrero & Mendoza, 2019; Bickenbach *et al.*, 2016).

En el Perú, el territorio está administrado en departamentos, provincias y distritos, siendo los distritos las unidades administrativas más pequeñas. No obstante, el mayor nivel de desagregación de series económicas de las que se dispone es a nivel de departamentos; aunque se vienen realizando esfuerzos para generar indicadores económicos a nivel distrital. Por un lado, el Instituto Nacional de Estadística e Informática (INEI) viene utilizando la metodología de Elbers, Lanjouw y Lanjouw (2003) para generar mapas de pobreza a nivel distrital. Por otro lado, el Programa de las Naciones Unidas para el Desarrollo (PNUD) viene realizando esfuerzos similares para generar el Índice de Desarrollo Humano (IDH) a nivel distrital. En un informe reciente, el PNUD (2019) brinda datos del IDH y sus componentes para las distintas divisiones administrativas del Perú (departamentos, provincias y distritos) para los años 2003, 2007, 2010, 2011, 2012, 2015, 2017, 2018 y 2019. No obstante, una de las principales desventajas es que los cálculos del IDH a nivel distrital utilizan encuestas de hogares, las cuales solo tienen representatividad a nivel departamental, por lo que podrían existir deficiencias en la actualización del IDH a nivel distrital. Por ello, el objetivo de esta investigación es ofrecer una metodología alternativa para estimar el PIB a nivel subnacional combinando información de censos y encuestas, estadísticas de cuentas nacionales y datos satelitales de luminosidad. Usando esta nueva metodología, hemos estimado el PIB a nivel de subdivisiones administrativas territoriales del Perú que no cuentan con estadísticas de cuentas nacionales (provincias y distritos). En este documento utilizaremos exclusivamente indicadores de PIB real a nivel nacional y subnacional.

Esta investigación se ha organizado en seis secciones. En la segunda sección, brindamos los detalles de los datos y metodologías empleados para la estimación del PIB a nivel subnacional y análisis de desarrollo regional en el Perú. En la tercera sección, presentamos los resultados de la estimación del PIB a nivel de provincias y distritos. En cuarta sección, brindamos análisis de las tendencias recientes de desarrollo regional en el Perú. En la quinta sección, presentamos las conclusiones. Por último, en la sexta sección, discutimos las limitaciones y futuras líneas de investigación.

2. Datos y metodología

2.1 Datos

Para obtener resultados consistentes con los hechos estilizados de las regiones de análisis, los cuales han sido corroborados *in situ* para varias regiones, ha sido necesario utilizar información de cuatro variables: (1) densidad poblacional, (2) ingresos de los hogares, (3) datos del PIB por departamento y (4) datos satelitales de luminosidad. En primer lugar, los datos de densidad poblacional los hemos generado utilizando información de los censos de población. En segundo lugar, la información de los ingresos de los hogares se ha obtenido del PNUD. En tercer lugar, los datos de PIB por departamento se han obtenido del INEI y de Seminario, Zegarra y Palomino (2019). En cuarto lugar, los datos satelitales de luminosidad se han obtenido de National Geophysical Data Center (NGDC) – National Oceanic and Atmospheric Administration's (NOAA).

2.1.1 Datos de censos y encuestas

Con base en los censos de población de 1993, 2007 y 2017, hemos generado series de población para cada una de las regiones de análisis para el período 1993-2018. Los datos censales se han obtenido del INEI. Las series se han generado usando la función Piecewise Cubic Hermite Interpolating Polynomial (PCHIP) implementada en Matlab. Asimismo, utilizando información del INEI, el Instituto Cuánto y ArcGIS, hemos calculado la extensión territorial en kilómetros cuadrados de cada una de las regiones de análisis. Combinando las series de población y la extensión territorial, hemos generado series de densidad poblacional a nivel de provincias y distritos para el período 1993-2018.

Utilizando el componente de ingresos del IDH calculado por el PNUD y el método de Geary y Stark (2002), generamos estadísticas de PIB a nivel de provincias y distritos para el año base 2007. De no contar con estadísticas de ingresos a nivel de subdivisiones territoriales administrativas, es necesario reproducir la metodología de Elbers, Lanjouw y Lanjouw (2003) o del PNUD para generar estadísticas de ingresos que sean representativas a nivel de las subdivisiones territoriales que queremos analizar.

2.1.2 Datos de cuentas nacionales

El INEI presenta series de PIB por departamento desde 1970; sin embargo, la composición del PIB regional desde 1970 hasta 1995 se encuentra en soles de 1979, mientras que la información desde 1995 hasta 2017, en soles de 1994. Desde el año 2007, el INEI presenta series de PIB por sector en soles de 2007. No obstante, Seminario *et al.* (2019), utilizando una metodología novedosa, reconstruyeron el PIB por departamento para el período 1795-2018 expresados en dólares de Geary-Khamis de 1990. Por ello, las series de PIB por departamento para el período 1993-2018 las obtenemos de Seminario *et al.* (2019).

2.1.3 Datos satelitales de luminosidad

Los datos satelitales de luminosidad tienen dos bloques: (1) 1992-2013 y (2) 2012-2018. Los datos del primer bloque tienen frecuencia anual, mientras que los datos del segundo bloque tienen frecuencia mensual.

Por un lado, los datos del período 1992-2013 fueron recolectados por el programa satelital Defense Meteorological Satellite Program Operational Lines-can System (DMSP-OLS) y mantenidos y procesados por el NGDC. Los satélites orbitan la Tierra y envían imágenes de cada ubicación entre 65 grados de latitud sur y 65 grados de latitud norte a una resolución de 30 segundos de arco (aproximadamente 1 km cuadrado en el ecuador) desde las 20:30 hasta las 22:00 hora local. Las imágenes se procesan para eliminar la cubierta de nubes, la nieve y las luces de corta duración (como los incendios forestales). A cada píxel (0,008241 grados, equivalentes a 0,9174 kilómetros cuadrados) en los datos de luminosidad se le asigna un número digital (ND) que representa su nivel de luminosidad¹. Los ND son enteros que van de 0 a 63.

¹ Consideramos que un arco de grado equivale a 111,32 kilómetros.

Por otro lado, los datos correspondientes al período 2012-2018 fueron generados por el Earth Observations Group (EOG) en NOAA – National Centers for Environmental Information (NCEI). Estos datos se obtienen utilizando datos de la Day/Night Band (DNB) del instrumento Visible Infrared Imaging Radiometer Suite (VIIRS) del satélite Suomi-NPP. Las imágenes satelitales de VIIRS abarcan todo el globo terrestre desde 75 grados de latitud norte hasta 65 grados de latitud sur. Estas imágenes se producen en cuadrículas geográficas a una resolución de 15 segundos de arco. A cada píxel (0,000007 grados, equivalentes a 0,000779 kilómetros cuadrados) de los datos de luminosidad se le asigna un ND que representa su nivel de luminosidad. Sin embargo, los ND de VIIRS tiene gran variabilidad y toman valores positivos y negativos con decimales, a diferencia de los ND del DMSP, que solo toman enteros positivos de 0 a 63.

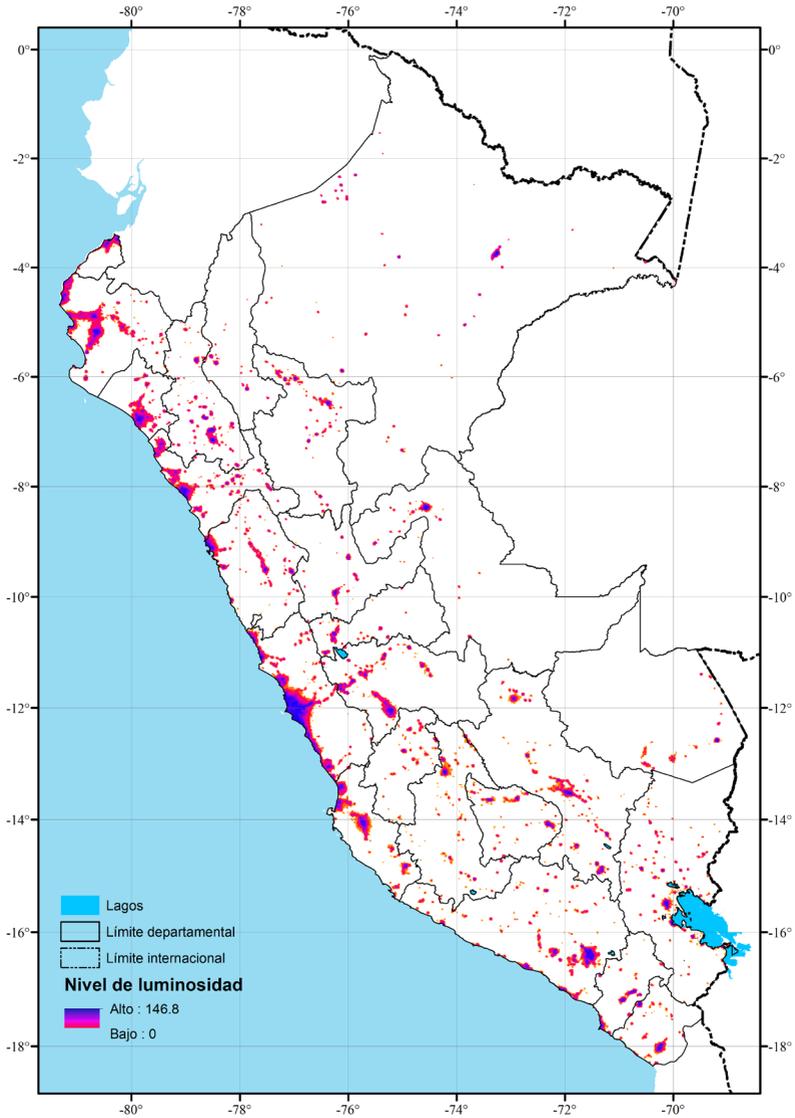
A diferencia de la literatura sobre luces nocturnas (Pinkovskiy & Sala-i-Martin, 2018; Michalopoulos & Papaioannou, 2014, 2018; Henderson *et al.*, 2012; Chen & Nordhaus, 2011) que suma los ND para construir índices de luminosidad, nuestra variable de luminosidad corresponde al promedio de los ND de los píxeles que componen la región geográfica de análisis. Sin embargo, utilizamos diferentes procedimientos porque los datos del DMSP son anuales, mientras que los de la VIIRS son mensuales. Existen grandes diferencias entre los mapas del DMSP (1992-2013) y VIIRS (04/2012-actualidad). Los datos de la VIIRS son muy superiores en sensibilidad y calibración respecto a los datos del DMSP (Zhao *et al.*, 2019).

El nivel de luminosidad anual para el período 1993-2013 lo calculamos usando la siguiente fórmula:

$$LA_{i,t} = \frac{\sum_{j=1}^{N_{i,t}} ND_{j,i,t}}{N_{i,t}}$$

donde $LA_{i,t}$ corresponde al nivel de luminosidad de la región i en el año t , $ND_{j,i,t}$ representa el ND del píxel j de la región en el año t , y $N_{i,t}$ corresponde al total de píxeles que tiene la región i en el año t .

Figura 1
Luminosidad a nivel departamental: 12/2017



Fuente: NGDC. Elaboración propia.

El nivel de luminosidad anual para el período 2013-2018 lo calculamos utilizando la siguiente fórmula:

$$LA_{i,t} = \frac{\sum_{m=1}^{12} LM_{i,m,t}}{12}$$

donde $LA_{i,t}$ corresponde al nivel de luminosidad de la región i en el año t ; y $LM_{i,m,t}$ corresponde al nivel de luminosidad de la región i en el mes m del año t y se calcula utilizando la siguiente fórmula:

$$LM_{i,m,t} = \frac{\sum_{j=1}^{N_{i,t}} ND_{j,i,m,t}}{N_{i,m,t}}$$

donde $ND_{j,i,m,t}$ representa el ND del píxel j de la región i en el mes m del año t , y $N_{i,m,t}$ corresponde al total de píxeles que tiene la región i en el mes m del año t . Los ND de los píxeles de luminosidad para el período 2013-2018 toman valores negativos y positivos. Por ejemplo, usando *grids* (cuadrados) de $(1/5) \times (1/5)$ de legua marina, encontramos que los ND de luminosidad del Perú en el año 2018 oscilan entre -0,05 y 146,8². En la figura 1, mostramos los *grids* de luminosidad con ND mayor de 0 para el año 2018 en el Perú.

Utilizando ArcGIS con Python, procesamos los datos de luminosidad a nivel de 24 departamentos, 196 provincias y 1.834 distritos. Sin embargo, cuando procesamos los datos a nivel de unidades administrativas pequeñas, encontramos presencia de muchos ceros, a los que podemos llamar *missing values* (MV). La presencia de MV puede ser consecuencia de varios factores: (1) el satélite no pasó por esa región, (2) el satélite pasó cuando el sistema de iluminación eléctrica se encontraba apagado debido a eventos climáticos, (3) la región no cuenta con un sistema de iluminación eléctrica. El procesamiento para el período 1992-2013 arroja 40 provincias y 1.022 distritos con MV; mientras que, para el período 2013-2018, arroja 4 provincias y 4 distritos con MV. La estrategia de corrección de MV consiste en empezar por las regiones más grandes y terminar por las regiones más pequeñas, es decir, primero corregimos las provincias y luego corregimos los distritos. Para corregir los MV, consideramos cuatro pasos: (1) obtener la luminosidad promedio de las regiones que se encuentran alrededor de la región que tiene MV, (2) obtener el nivel de luminosidad de la región

² La legua marina mide 5,555 km (5.555 m) (equivalentes a 1/20 de un grado de meridiano terrestre). Se utilizó la legua marina porque Seminario *et al.* (2019) encontraron que es una medida eficiente para generar el área poblada en el Perú.

con MV utilizando las tasas de crecimiento de la luminosidad promedio de la región de mayor tamaño a la que pertenece la región de menor tamaño, (3) obtener el nivel de luminosidad utilizando la tasa de crecimiento del nivel de luminosidad promedio de los distritos vecinos filtrada con una ventana de Hann de tres observaciones, (4) elegir el nivel de luminosidad promedio mínimo de (2) o (3) y asignarlo a la región que tiene MV³.

Después de corregir los MV, ajustamos los datos de la VIIRS a los datos del DMSP utilizando la siguiente fórmula:

$$L_{i,t} = LA_{i,t_0} * \frac{LBC_{i,t}}{LBC_{i,t_0}}$$

donde $L_{i,t}$ corresponde al nivel de luminosidad final de la región i en el año t ; LA_{i,t_0} representa el nivel de luminosidad la región i en el año t_0 , el cual corresponde al año en el que se tiene datos del DMSP y también se tiene datos de la VIIRS (es decir, corresponde al año 2013); $LBC = (LB + \theta) > 0$ corresponde al nivel de luminosidad de la VIIRS ajustado al nivel de luminosidad del DMSP para el período 2013-2018; y θ es un valor constante positivo necesario para eliminar los valores negativos del nivel de luminosidad calculado con las imágenes satelitales provenientes de la VIIRS⁴.

2.2 Metodología

Para generar series de PIB de regiones territoriales pequeñas que no cuentan con estadísticas de cuentas nacionales, ha sido necesario describir la metodología en tres subsecciones. En la primera subsección, describimos el Método de Geary y Stark (2002). En la segunda subsección, describimos las etapas de la metodología utilizada para estimar el PIB. En la tercera subsección, describimos la metodología utilizada para el análisis espacial de las series del PIB a nivel de provincias y distritos.

2.2.1 Método de Geary-Stark

En un ensayo, Geary y Stark (2002) desarrollaron un método que permitía distribuir el PIB del Reino Unido entre sus principales regiones: Escocia,

³ La ventana de Hann fue inventada por el meteorólogo Julius von Hann y, a partir de su aparición en la publicación *The measurement of power spectra, from the point of view of Communications Engineering* del estadista John Tukey, en 1959, la aplicación de esta ventana es denominada «hanning».

⁴ En el Perú, para corregir los valores negativos del nivel de luminosidad de la VIIRS para el período 2013-2018, consideramos θ igual a 10.

Gales, Inglaterra e Irlanda. El método requiere conocer el empleo y los salarios de las distintas regiones. Se asume que la razón que existe entre el PIB per cápita y una región y el promedio nacional es proporcional a la que hay entre los salarios regionales y los salarios promedio nacionales. Posteriormente, diversos autores han utilizado el método propuesto por Geary y Stark para derivar estadísticas de PIB a nivel subregional en distintos países (Badia-Miró, 2015; Crafts, 2004).

En el Perú, no existe información requerida por el método de Geary & Stark; por ello, Seminario *et al.* (2019) consideraron necesario modificar el método para obtener un PIB departamental consistente con los hechos estilizados para el período 1795-2018. Por un lado, los autores dividen la estimación de las series departamentales en tres fases: (1) 1896-1929, (2) 1929-1970 y (3) 1970-2018. Por otro lado, los estimados del PIB departamental registran la suma de la producción distribuida en cuatro grandes sectores: (1) agricultura (interna y de exportación), (2) minería, (3) servicios gubernamentales y (4) otros sectores (incluye industria).

En línea con Seminario *et al.* (2019), adaptamos el método de Geary y Stark para derivar series de PIB a nivel de provincias y distritos en el Perú para el período 1993-2018. Por un lado, utilizamos el método modificado y las series de ingresos de los hogares para generar series de PIB a nivel de provincias y distritos para el año base 2007. Por otro lado, utilizamos el método modificado para ajustar las series de PIB a niveles de provincias/distritos a las series de PIB a nivel de departamentos/provincias.

2.2.2 Estimación del PIB

La estrategia para estimar el PIB de regiones administrativas pequeñas que no cuentan con estadísticas de cuentas nacionales consiste en estimar el PIB de regiones administrativas grandes y luego utilizar estos datos para estimar el PIB de regiones administrativas pequeñas. En el Perú, contamos con estadísticas de actividad económica a nivel departamental, pero no a nivel de provincias o distritos; por ello, para estimar el PIB a nivel distrital, primero debemos estimar el PIB a nivel provincial ajustado al PIB de los departamentos.

La estimación del PIB de regiones territoriales pequeñas consta de cuatro etapas: (1) generar series de luminosidad y densidad poblacional, (2) estimar el PIB de un año base, (3) estimar los parámetros de una función de producción, (4) generar series de PIB para todo el período de análisis, (5) ajustar las series de PIB estimadas de las regiones territoriales menores al PIB de las regiones territoriales mayores.

En primer lugar, la metodología para la generación de series del nivel de luminosidad y densidad poblacional la describimos en la sección 2.1.

En segundo lugar, generamos series de PIB a nivel de provincias/distritos utilizando el método adaptado de Geary y Stark y la información de ingresos de los hogares derivadas por el PNUD con la siguiente fórmula:

$$y_{i,k,t_0} = y_{k,t_0} * \frac{I_{i,k,t_0}}{\sum_j^N I_{j,k,t_0}}$$

donde $y_{i,k}$ corresponde al PIB estimado con datos de ingresos de la región territorial menor i que pertenece a la región territorial mayor k ; y_k representa el PIB de la región territorial mayor k ; $I_{i,k}$ corresponde al ingreso de la región territorial menor i que pertenece a la región territorial mayor k ; $\sum_j^N I_{j,k,t_0}$ representa el ingreso total de todas las regiones territoriales menores que pertenecen a la región territorial mayor k ; N corresponde a la cantidad total de regiones territoriales menores que posee la región territorial mayor k ; y t_0 corresponde al año base.

En tercer lugar, para estimar los parámetros de una función de producción fue necesario encontrar una función adecuada para el uso correcto de las series del nivel de luminosidad y densidad poblacional. Realizando diferentes pruebas, encontramos que el nivel de luminosidad como *proxy* de capital y la densidad poblacional como *proxy* de fuerza laboral reproducen las series de PIB a nivel de departamentos que provienen de cuentas nacionales usando una función de producción cúbica. Por tanto, estimamos los parámetros de la función de producción cúbica para las regiones que no cuentan con estadísticas de cuentas nacionales (provincias / distritos) utilizando la siguiente especificación econométrica:

$$\ln(y_{i,t_0}) = \alpha_0 + \alpha_1 \ln(L_{i,t_0}) + \alpha_2 \ln(L_{i,t_0})^2 + \alpha_3 \ln(L_{i,t_0})^3 + \alpha_4 \ln(D_{i,t_0}) + \mu_{i,t_0}$$

donde y_i representa el PIB de la región i , L_i corresponde a nivel de luminosidad promedio de la región i , D_i representa a la densidad poblacional de la región i , y t_0 corresponde al año base.

En cuarto lugar, utilizando los coeficientes estimados en el paso anterior, estimamos el PIB a nivel de provincias/distritos usando la siguiente fórmula:

$$\ln(\hat{y}_{i,k,t}) = \hat{\alpha}_0 + \hat{\alpha}_1 \ln(L_{i,k,t}) + \hat{\alpha}_2 \ln(L_{i,k,t})^2 + \hat{\alpha}_3 \ln(L_{i,k,t})^3 + \hat{\alpha}_4 \ln(D_{i,k,t})$$

donde $(\hat{y}_{i,k,t})$, $L_{i,k,t}$ y $D_{i,k,t}$ corresponden al PIB estimado, nivel de luminosidad promedio y densidad poblacional de la región territorial menor que pertenece a la región territorial mayor k en el período t , respectivamente. Asimismo, dado

que las series estimadas de PIB están expresadas en logaritmos, procedemos a obtener las series de PIB en unidades del año base utilizando la siguiente fórmula:

$$y_{i,k,t}^E = y_{i,k,t_0} * e^{[\ln(\hat{y}_{i,k,t}) - \ln(\hat{y}_{i,k,t_0})]}$$

donde $y_{i,k,t}^E$ corresponde al PIB estimado de la región territorial menor i que pertenece a la región territorial mayor k en el período t , y_{i,k,t_0} corresponde al PIB estimado con datos de ingresos de la región territorial menor i que pertenece a la región territorial mayor k en el año base, $\ln(\hat{y}_{i,k,t})$ corresponde al PIB estimado con datos del nivel de luminosidad y densidad poblacional de la región territorial menor i que pertenece a la región territorial mayor k en el período t , y $\ln(\hat{y}_{i,k,t_0})$ corresponde al PIB estimado con datos del nivel de luminosidad y densidad poblacional de la región territorial menor que pertenece a la región territorial mayor k en el año base.

En quinto lugar, utilizando la versión modificada del método de Geary y Stark, ajustamos los datos del PIB de regiones territoriales menores con datos del PIB de regiones territoriales mayores. Por un lado, usamos las series del PIB a nivel de departamentos (regiones territoriales mayores) obtenidas del INEI para ajustar las series estimadas del PIB a nivel de provincias (regiones territoriales menores). Por otro lado, utilizamos las series del PIB de provincias (regiones territoriales mayores) para ajustar las series del PIB estimadas a nivel de distritos (regiones territoriales menores). En concreto, la estimación del PIB de regiones territoriales menores requiere de series del PIB de regiones territoriales mayores. Las series estimadas del PIB con datos satelitales de luminosidad y densidad poblacional las ajustamos utilizando la siguiente fórmula:

$$y_{i,k,t}^F = y_{k,t}^F * \frac{y_{i,k,t}^E}{\sum_j^N y_{j,k,t}^E}$$

donde $y_{i,k,t}^F$ corresponde al PIB final de la región territorial menor i que pertenece a la región territorial mayor k en el período t , $y_{k,t}^F$ representa el PIB final de la región territorial mayor k en el período t , $y_{i,k,t}^E$ corresponde al PIB estimado de la región territorial menor que pertenece a la región territorial mayor k en el período t , $\sum_j^N y_{j,k,t}^E$ representa el PIB total estimado de todas las regiones territoriales menores que pertenecen a la región territorial mayor k en el período t , y N corresponde a la cantidad total de regiones territoriales menores que posee la región territorial mayor k .

2.2.3 Generación de grupos de regiones

Utilizando datos a nivel de departamentos, Seminario *et al.* (2019) encontraron que la distribución del PIB per cápita del Perú es bimodal. Por ello, consideramos necesario identificar metodologías que permitan determinar umbrales óptimos para clasificar las regiones según alguna variable de interés.

La elección de umbrales óptimos de los datos es idónea para elaborar mapas georreferenciados que permitan verificar la consistencia de las series generadas con los hechos estilizados observados en diferentes regiones. En esta investigación, exploramos cuatro métodos de clasificación de regiones: (1) cuantiles, (2) cortes naturales (Jenks), (3) Otsu y (4) Atlas del Banco Mundial.

En primer lugar, los cuantiles son aquellos valores de la variable de interés que, ordenados de menor a mayor, dividen a la distribución de los datos en partes, de tal manera que cada una de ellas contiene el mismo número de frecuencias.

En segundo lugar, los cortes naturales se basan en el algoritmo de optimización de Jenks, diseñado para determinar la mejor disposición de los valores en diferentes clases. Esto se hace buscando minimizar la desviación promedio de cada clase de la media de la clase, mientras se maximiza la desviación de cada clase de las medias de los otros grupos. En otras palabras, el método busca reducir la varianza dentro de las clases y maximizar la varianza entre clases.

En tercer lugar, el método de Otsu (1979) calcula el valor umbral de forma que la dispersión dentro de cada segmento sea lo más pequeña posible, pero, al mismo tiempo, la dispersión sea lo más alta posible entre segmentos diferentes.

En cuarto lugar, el método Atlas⁵ del Banco Mundial divide las economías según sus niveles de ingreso nacional bruto per cápita (INB pc) en cuatro grupos: (1) bajo, (2) medio bajo, (3) medio alto y (4) alto. Dado que las series de PIB que generamos están en dólares Geary-Khamis de 1990, ajustamos los umbrales de INB pc del Banco Mundial en 1990 a umbrales de PIB per cápita y PIB por km² utilizando las siguientes fórmulas:

$$y_{C,t_0}^{BM} = g_{t_0} * x_{U,t_0}^{BM}$$

$$y_{U,t_0}^N = f_{t_0} * y_{U,t_0}^{BM}$$

donde x_{U,t_0}^{BM} corresponde al umbral U del INB pc derivado por el Banco Mundial en t_0 (año 1990), y_{U,t_0}^{BM} corresponde al umbral U del PIB per cápita o PIB por

⁵ <https://datahelpdesk.worldbank.org/knowledgebase/articles/378832-what-is-the-world-bank-atlas-method>

km² ajustado con series nacionales del Banco Mundial en t_0 , y_{i,t_0}^N corresponde al umbral U del PIB per cápita o PIB por km² ajustado con series nacionales derivadas por Seminario *et al.* (2019) en t_0 , g_{t_0} corresponde al factor de ajuste del INB pc al PIB per cápita o PIB por km² a nivel nacional derivadas por el Banco Mundial en t_0 , y f_{t_0} corresponde al factor de ajuste del PIB per cápita o PIB por km² derivados por el Banco Mundial a las estadísticas derivadas por Seminario *et al.* (2019) a nivel nacional en t_0 .

Realizando diferentes pruebas, encontramos que el método de Otsu es el más idóneo para determinar umbrales de los datos en presencia de multimodalidad. Por ello, utilizamos los umbrales de Otsu para generar grupos de regiones y elaborar mapas georreferenciados de PIB. No obstante, para realizar un análisis comparativo con la metodología del Banco Mundial, elaboramos mapas georreferenciados utilizando los umbrales de PIB per cápita y PIB por km² generados con base en los umbrales de INB pc derivados por el Banco Mundial.

2.2.4 Desigualdad regional

Usando el índice de Gini y las estadísticas de PIB pc y PIB por km², analizamos la evolución de la desigualdad regional a nivel de provincias y distritos. De acuerdo con Theil (1967), el índice de Gini está definido como:

$$G = \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n |x_i y_j - x_j y_i|$$

donde n es la cantidad de regiones, y_i es la participación individual de la región i en la población total (o área total), y x_i es el PIB pc (o PIB por km²). Por tanto, G es la mitad de un promedio ponderado de todas las diferencias absolutas entre los PIB pc (o PIB por km²). Los pesos están definidos por la multiplicación de las respectivas participaciones de la población (o área) respecto al total. El mínimo valor del índice es cero y se alcanza cuando el PIB pc (o PIB por km²) de todas las regiones es igual.

2.2.5 Desarrollo regional

De acuerdo con Ivanović (1963), un indicador de desarrollo regional se puede calcular utilizando la siguiente fórmula:

$$D_{r,t} = \sum_{i=1}^n \frac{|d_{i,r,t}|}{\sigma_{i,t}} \prod_{j=1}^{i-1} (1 - r_{ij,t})$$

donde $D_{r,t}$ corresponde el nivel de desarrollo de la región r en el período t , $d_{i,r,t}$ corresponde a la distancia de la variable i respecto a su valor mínimo de la región r en el período t ($d_{i,r,t} = x_{i,r,t} - x_{i,min,t}$), $r_{i,j,t}$ corresponde a la correlación entre la variable i y la variable j en el período t , σ_i corresponde a la desviación estándar de la variable i en el período t , y n corresponde al número de variables consideradas para calcular el indicador de desarrollo.

Siguiendo la metodología de Ivanović (1963) y las estadísticas de PIB pc, densidad poblacional y el nivel de urbanización, generamos un nuevo indicador de desarrollo regional a nivel de provincias y distritos. El nivel de urbanización lo calculamos dividiendo la población urbana entre la población total. La población urbana corresponde a la población que vive en áreas urbanas.

De acuerdo con INEI (2018), existen dos criterios para delimitar la frontera entre lo urbano y lo rural. El primer criterio es para fines operativos de los censos de población, donde se considera como ámbito urbano a las áreas con un mínimo de 100 viviendas agrupadas contiguamente (en promedio, 500 habitantes) y a todos los centros poblados capitales de distritos (aun cuando no reúnan la condición indicada). El segundo criterio es tomado en cuenta en las encuestas de hogares, donde se considera ámbito urbano a las áreas de los centros poblados con más 2.000 habitantes donde sus viviendas se encuentran agrupadas de manera contigua, formando manzanas y calles.

3. Estimación del PIB a nivel subnacional

En esta sección, presentamos los resultados de las estimaciones de la función cúbica a nivel de provincias y distritos, regiones que no cuentan con series de PIB basadas en cuentas nacionales. Asimismo, presentamos mapas georreferenciados de los resultados, los cuales nos permiten analizar la consistencia de los datos generados con los hechos estilizados.

En las regresiones, utilizamos PIB, PIB per cápita y PIB por km² con el objetivo de seleccionar los parámetros que ajustan mejor los datos. Las regresiones han sido estimadas con datos del PIB para el año base 2007. Los datos del PIB de 2007 se han obtenido utilizando el método de Geary y Stark (2002) e información de ingresos de los hogares a nivel de provincias y distritos calculados por el PNUD.

3.1 Estimación del PIB a nivel de distritos

Las estadísticas representadas en la figura 2 sugieren una relación positiva entre el nivel de luminosidad, la densidad poblacional y el PIB por km² en el año 2007. De acuerdo con estas figuras, el incremento de un 1% en el nivel de luminosidad y en la densidad poblacional incrementa el PIB por km² en un 0,74 y un 1,11% respectivamente. Las estadísticas de estas figuras también sugieren que la densidad poblacional y el nivel de luminosidad explican alrededor del 93 y el 54% de la variabilidad del PIB por km² respectivamente. De acuerdo con estos resultados, podemos indicar que el nivel de luminosidad y la densidad poblacional son variables claves para explicar la dinámica del PIB por km².

La tabla 1 muestra los resultados de la estimación de la función de producción cúbica logarítmica. Esta función de producción tiene dos factores de producción: (1) nivel de luminosidad como *proxy* de capital y (2) la densidad

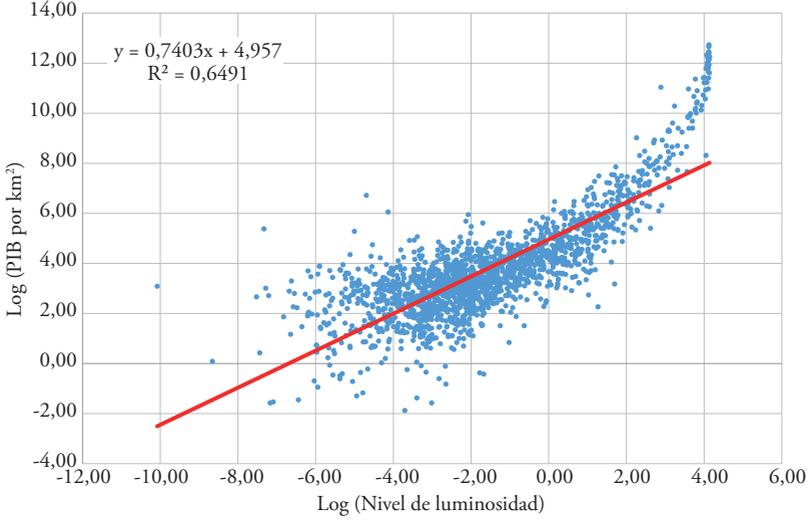
poblacional como *proxy* de fuerza laboral. Los datos de estas regresiones son a nivel de 1.834 distritos del Perú y corresponden al año base 2007.

La tabla 1.A sugiere una relación positiva entre el nivel de luminosidad, la densidad poblacional y el PIB a nivel distrital. De acuerdo con los resultados, el incremento del 1% en el nivel de luminosidad aumenta el nivel del PIB en un 0,43% cuando se considera nivel de luminosidad de grado uno (modelo 1), en un 0,20% cuando se considera el nivel de luminosidad y la densidad poblacional de grado uno (modelo 2) y en un 0,38% cuando se considera el nivel de luminosidad en una función cúbica (modelo 3). Los resultados de los modelos 2 y 3 también indican que existe una relación positiva entre el PIB y la densidad poblacional. El modelo 3 sugiere que el incremento del 1% en la densidad poblacional incrementa el PIB en un 0,23%. Los resultados de esta tabla indican que el nivel de luminosidad y densidad de población en una función cúbica explican alrededor del 49% de la varianza del PIB.

La tabla 1.B sugiere una relación positiva del PIB per cápita con el nivel de luminosidad y una relación negativa con la densidad poblacional a nivel distrital. Los resultados del modelo 3 indican que la elasticidad es del 0,29% cuando el nivel de luminosidad es de grado uno, del 0,05% cuando el nivel de luminosidad es de grado dos y del 0,003% cuando el nivel de luminosidad es de grado tres. Los resultados indican que la elasticidad de densidad poblacional es de un -0,05% cuando solo interactúa con el nivel de luminosidad de grado uno y de un -0,15% cuando interactúa con los niveles de luminosidad de grado uno, dos y tres.

Figura 2
Relación entre PIB por km², nivel de luminosidad y densidad poblacional
(estadísticas a nivel distrital para el año 2007)

A: Relación entre PIB por km² y nivel de luminosidad



B: Relación entre PIB por km² y densidad poblacional

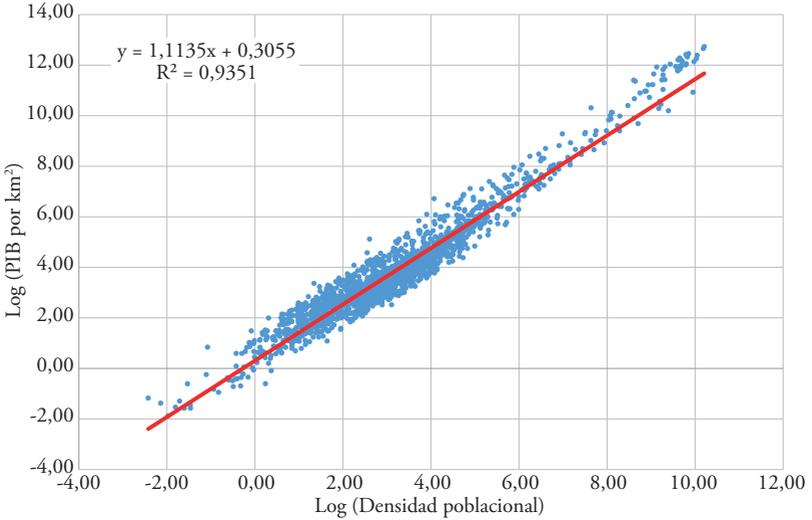


Tabla 1
Estimación de una función cúbica de producción a nivel de distritos

A: Log (PIB)

	(1)	(2)	(3)
Log (Luminosidad promedio)	0,4264*** (0,0128)	0,2024*** (0,0180)	0,3813*** (0,0252)
Log (Densidad poblacional)		0,3743*** (0,0226)	0,2251*** (0,0257)
Log (Luminosidad promedio) ²			0,0586*** (0,0065)
Log (Luminosidad promedio) ³			0,0025** (0,0010)
R ²	0,3779	0,4590	0,4943
R ² -Ajustado	0,3776	0,4584	0,4932
F-Estadística	1.113,0381	776,7870	447,0077

B: Log (PIB per cápita)

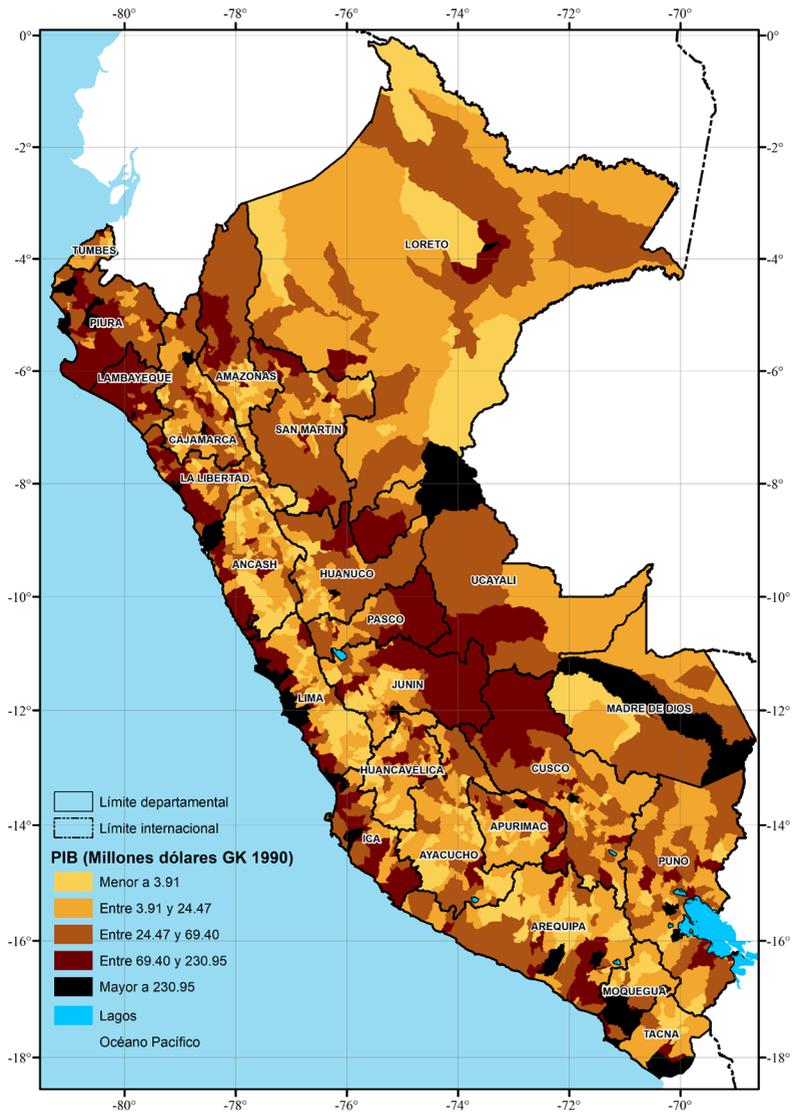
	(1)	(2)	(3)
Log (Luminosidad promedio)	0,1418*** (0,0048)	0,1688*** (0,0073)	0,2853*** (0,0094)
Log (Densidad poblacional)		-0,0452*** (0,0091)	-0,1525*** (0,0096)
Log (Luminosidad promedio) ²			0,0449*** (0,0024)
Log (Luminosidad promedio) ³			0,0028*** (0,0004)
R ²	0,3193	0,3283	0,4629
R ² -Ajustado	0,3189	0,3275	0,4617
F-Estadística	859,2176	447,4239	394,1195

C: Log (PIB por km²)

	(1)	(2)	(3)
Log (Luminosidad promedio)	0,7403*** (0,0127)	0,1688*** (0,0073)	0,2853*** (0,0094)
Log (Densidad poblacional)		0,9548*** (0,0091)	0,8475*** (0,0096)
Log (Luminosidad promedio) ²			0,0449*** (0,0024)
Log (Luminosidad promedio) ³			0,0028*** (0,0004)
R ²	0,6491	0,9499	0,9599
R ² -Ajustado	0,6489	0,9498	0,9598
F-Estadística	3.389,4097	17.355,3885	10.955,9168

Notas. Esta tabla presenta las estimaciones de funciones de producción logarítmica que considera el nivel de luminosidad como *proxy* de capital y la densidad poblacional como *proxy* de fuerza laboral en el Perú. La función de producción se ha estimado utilizando mínimos cuadrados generalizados. Las estimaciones consideran datos para una muestra de 1.834 distritos en el año base 2007. Las estimaciones de las columnas (1), (2) y (3) consideran como factor de producción al nivel de luminosidad de grado uno, el nivel de luminosidad y densidad poblacional de grado uno, y el nivel de luminosidad y densidad poblacional en una función cúbica, respectivamente. Las estimaciones de los paneles A, B y C consideran como variable de resultado el PIB, el PIB per cápita y el PIB por km², respectivamente. Los errores estándar robustos se encuentran entre paréntesis. La significancia estadística al 1%, 5% y 10% se indica con ***, ** y *, respectivamente.

Figura 3
PIB a nivel de distritos, 2018



Fuete: elaboración propia.

La tabla 1.C sugiere una relación positiva entre el nivel de luminosidad, la densidad poblacional y el PIB por km² a nivel distrital. Los resultados del modelo 3 indican que el aumento del 1% en la población por km² aumenta el PIB por km² en un 0,72%; mientras que el aumento del 1% en el nivel de luminosidad de grado uno, dos y tres incrementa el PIB por km² en un 0,28, un 0,05 y un 0,03%, respectivamente.

De acuerdo con los resultados de la tabla 1, el R² alcanza un valor del 54,59% cuando la variable dependiente es el PIB, un 53,79% cuando la variable dependiente es el PIB per cápita y un 95,04% cuando la variable dependiente es el PIB por km². Por tanto, podemos indicar que el nivel de luminosidad y la densidad poblacional son más útiles para explicar el comportamiento del PIB por km². Observamos el mismo patrón si analizamos el R² ajustado; por lo tanto, podemos confiar en los resultados del R².

Utilizando los resultados del modelo 3 de la tabla 1.C, estimamos el PIB a nivel de 1.834 distritos del Perú con la metodología descrita en la sección 2.2.2. Asimismo, utilizando la metodología descrita en la sección 2.2.3, agrupamos los distritos en cinco clases según su PIB, considerando cortes variables para cada año.

En la figura 3, mostramos la distribución regional del PIB a nivel distrital. En estas figuras, el color más oscuro corresponde a los distritos que pertenecen a la cola superior de la distribución, mientras que el color más claro corresponde a los distritos que pertenecen a la cola inferior de la distribución. Es decir, el color más oscuro corresponde a los distritos que pertenecen a la clase de mayores niveles de PIB.

En la figura 3, mostramos la agrupación de los distritos según su nivel de PIB en cinco clases para 2018. Estas figuras también sugieren que gran parte de los distritos que han generado mayor nivel de PIB se localizan en la Costa, algunos en la Sierra y muy pocos en la Selva. Los resultados indican que la distribución espacial de los distritos con mayor nivel de PIB depende de la región de análisis. En la Costa, el departamento de Lima concentra la mayor cantidad de distritos que generan altos niveles de PIB. En la Sierra, los distritos que generan mayor nivel de PIB encuentran principalmente en las regiones mineras. En la Selva, gran parte de los distritos que generan mayor nivel de PIB se localizan en la Yunga Fluvial y la Selva Alta porque estas regiones son aptas para la agricultura de exportación; también hay distritos que generan altos niveles de PIB en la Selva Baja, pero se localizan en las capitales departamentales: Iquitos, Pucallpa y Puerto Maldonado.

Utilizando el PIB a nivel distrital y la población a nivel de CP del año 2007, estimamos el PIB a nivel de CP y *grids* de 1/5 de legua marina. Los detalles metodológicos de estas estimaciones se encuentran en los anexos 1 y 2.

3.2 Estimación del PIB a nivel de provincias

La tabla 2 muestra los resultados de la estimación a nivel de 196 provincias del Perú de la función de producción cúbica logarítmica cuyos factores de producción son el nivel de luminosidad y la densidad poblacional. Los datos de estas regresiones corresponden al año base 2007.

Tabla 2
Estimación de una función cúbica de producción a nivel de provincias

A: Log (PIB)

	(1)	(2)	(3)
Log (Luminosidad promedio)	0,5533*** (0,0412)	0,4376*** (0,0678)	0,7249*** (0,0893)
Log (Densidad poblacional)		0,1784** (0,0835)	0,0756 (0,0839)
Log (Luminosidad promedio) ²			0,0881*** (0,0267)
Log (Luminosidad promedio) ³			0,0042 (0,0045)
R ²	0,4814	0,4934	0,5459
R ² -Ajustado	0,4788	0,4882	0,5364
F-Estadística	180,1190	93,9919	57,4125

B: Log (PIB per cápita)

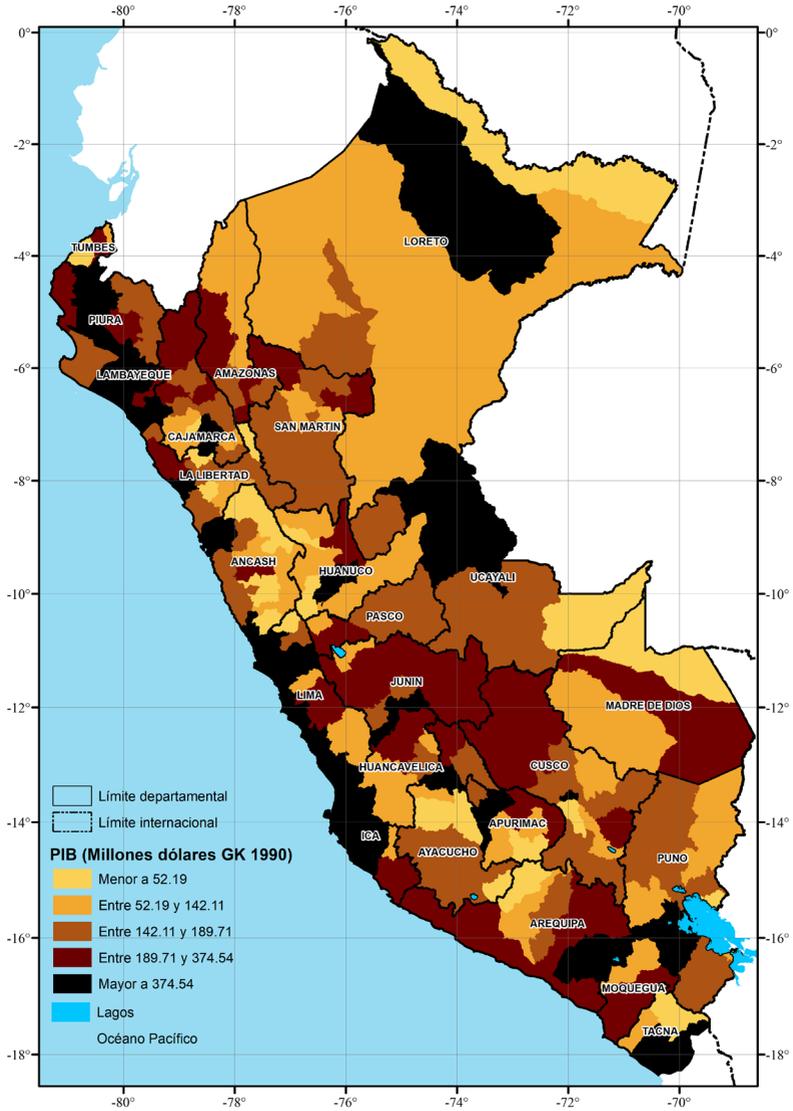
	(1)	(2)	(3)
Log (Luminosidad promedio)	0,1705*** (0,0182)	0,3145*** (0,0273)	0,4628*** (0,0345)
Log (Densidad poblacional)		-0,2221*** (0,0337)	-0,2796*** (0,0324)
Log (Luminosidad promedio) ²			0,0508*** (0,0103)
Log (Luminosidad promedio) ³			0,0033* (0,0017)
R ²	0,3121	0,4387	0,5379
R ² -Ajustado	0,3086	0,4329	0,5283
F-Estadística	88,0312	75,4193	55,5897

C: Log (PIB por km²)

	(1)	(2)	(3)
Log (Luminosidad promedio)	0,8186*** (0,0319)	0,3145*** (0,0273)	0,4628*** (0,0345)
Log (Densidad poblacional)		0,7779*** (0,0337)	0,7204*** (0,0324)
Log (Luminosidad promedio) ²			0,0508*** (0,0103)
Log (Luminosidad promedio) ³			0,0033* (0,0017)
R ²	0,7728	0,9397	0,9504
R ² -Ajustado	0,7717	0,9391	0,9493
F-Estadística	660,0543	1.503,4717	913,9834

Notas. Esta tabla presenta las estimaciones de funciones de producción logarítmica que considera el nivel de luminosidad como *proxy* de capital y la densidad poblacional como *proxy* de fuerza laboral en el Perú. La función de producción se ha estimado utilizando mínimos cuadrados generalizados. Las estimaciones consideran datos para una muestra de 196 provincias en el año base 2007. Las estimaciones de las columnas (1), (2) y (3) consideran como factor de producción al nivel de luminosidad de grado uno, el nivel de luminosidad y densidad poblacional de grado uno, y el nivel de luminosidad y densidad poblacional en una función cúbica, respectivamente. Las estimaciones de los paneles A, B y C consideran como variable de resultado el PIB, el PIB per cápita y el PIB por km², respectivamente. Los errores estándar robustos se encuentran entre paréntesis. La significancia estadística al 1%, 5% y 10% se indica con ***, ** y *, respectivamente.

Figura 4
PIB a nivel de provincias, 2018



Fuente: elaboración propia.

La tabla 2.A muestra la relación entre nivel de luminosidad, densidad poblacional y PIB. Los resultados sugieren una relación positiva y significativa a un nivel de confianza del 90% entre el PIB y el nivel de luminosidad. De acuerdo con los resultados, el incremento de un 1% en el nivel de luminosidad aumenta el nivel del PIB en un 0,55% cuando se considera el nivel de luminosidad de grado uno, en un 0,44% cuando se considera el nivel de luminosidad y la densidad poblacional de grado uno, y en un 0,72% cuando se considera la función de producción cúbica en el nivel de luminosidad. Los resultados también indican que existe una relación positiva entre el PIB y la densidad poblacional. Sin embargo, la relación positiva entre PIB y población por km² solo es significativa cuando no se considera el nivel de luminosidad de grado dos y grados tres.

La tabla 2.B muestra la relación entre nivel de luminosidad, densidad poblacional y PIB per cápita. Los resultados sugieren una relación positiva entre el PIB per cápita y el nivel de luminosidad de grado uno, dos y tres. Sin embargo, el valor de las elasticidades es menor a medida que el grado del nivel de luminosidad aumenta. Los resultados de la columna 2 indican que la elasticidad es del 0,46% cuando el nivel de luminosidad es de grado uno, del 0,05% cuando el nivel de luminosidad es de grado dos y del 0,003% cuando el nivel de luminosidad es de grado tres. Los resultados también sugieren una relación negativa entre la densidad poblacional a nivel de provincias. Los resultados indican que los efectos negativos de la densidad poblacional oscilan entre un 0,22% cuando solo interactúa con el nivel de luminosidad de grado uno y un 0,28% cuando interactúa con el nivel de luminosidad de grado uno, dos y tres.

La tabla 2.C muestra la relación entre nivel de luminosidad, densidad poblacional y PIB por km². Los resultados sugieren una relación positiva entre el PIB por km² y el nivel de luminosidad de grado uno, dos y tres. Los resultados también sugieren una relación positiva entre la densidad poblacional y el PIB por km². Los resultados indican que el aumento de un 1% en la población por km² aumenta el PIB por km² en un 0,72%. De acuerdo con los resultados, el R² alcanza un valor de un 54,59% cuando la variable dependiente es el PIB, un 53,79% cuando la variable dependiente es el PIB per cápita y un 95,04% cuando la variable dependiente es el PIB por km². Por tanto, podemos indicar que el nivel de luminosidad y la densidad poblacional son más útiles para explicar el comportamiento del PIB por km².

En línea con las estimaciones a nivel distrital, estimamos el PIB a nivel de 196 provincias del Perú utilizando los resultados del modelo 3 de la tabla 2.C. En la figura 4, mostramos la distribución espacial del PIB estimado a nivel provincial. De acuerdo con esta figura, la mayor cantidad de provincias que generan los mayores niveles de PIB se localizan en los departamentos con costa. En la figura también observamos que las provincias de los departamentos del centro con Yunga Fluvial y Selva Alta generan altos niveles de PIB. Los resultados también sugieren que los departamentos con Selva Baja tienen provincias que generan altos niveles de PIB. No obstante, las estadísticas a nivel de distritos sugieren que no todos los distritos de una provincia tienen el mismo desempeño. Por lo tanto, para evitar introducir sesgos en los análisis empíricos de economía regional, se recomienda considerar estadísticas para diferentes resoluciones espaciales.

4. Tendencias recientes de desarrollo regional

Utilizando las estadísticas generadas en esta investigación, analizamos las tendencias recientes de desarrollo regional a nivel de distritos. En primer lugar, analizamos la distribución espacial del PIB pc y PIB por km². En segundo lugar, analizamos la evolución de la desigualdad regional del PIB pc y PIB por km². En tercer lugar, analizamos el desarrollo regional utilizando un nuevo indicador, el cual combina PIB per cápita, densidad poblacional y urbanización.

4.1 Distribución espacial del PIB pc y PIB por km²

Adaptando la escala del Banco Mundial para clasificar los países (véase la sección 2.2.3 para los detalles de la metodología), clasificamos los distritos en cuatro grupos según sus niveles de PIB pc y PIB por km²: (1) bajo, (2) medio bajo, (3) medio alto y (4) alto. En la tabla, mostramos la cantidad de distritos en cada grupo entre 1993 y 2018. Los resultados de esta tabla indican que la mayor proporción de distritos pertenecen a la cola inferior de la distribución tanto del PIB pc como del PIB por km². Los resultados de esta tabla sugieren que la distribución de la cantidad de distritos es más uniforme en el PIB por km² que en el PIB pc. Por ejemplo, en 2018, el grupo de distritos de PIB pc medio bajo concentró la mayor cantidad de distritos (74,21%), mientras que el grupo de PIB por km² bajo concentró la mayor cantidad de distritos (39,1%). En las figuras 5 y 6, mostramos la distribución espacial de los distritos según sus niveles del PIB pc y PIB por km² para el año 2018. En estas figuras, el color más oscuro indica mayor nivel de PIB pc o PIB por km², mientras que el color más claro indica menor nivel de PIB pc o PIB por km².

Tabla 3
Cantidad de distritos según sus niveles de PIB pc y PIB por km²

	PIB pc				PIB por km ²			
	1993		2018		1993		2018	
	N	%	N	%	N	%	N	%
Bajo	1.177	64,18	286	15,59	1.071	58,40	717	39,09
Medio bajo	698	38,06	1361	74,21	486	26,50	609	33,21
Medio alto	19	1,04	183	9,98	125	6,82	227	12,38
Alto	0	0	4	0,22	152	8,29	281	15,32
Nacional	1.834							

Por un lado, la figura 5 sugiere que la mayoría de los distritos que tienen PIB pc medio alto se localizan en la Costa del centro y sur del país. Las estadísticas de esta figura también sugieren que algunos distritos de los departamentos de Junín, Huancavelica, Apurímac y la sierra de Lima pertenecen al grupo que tiene PIB pc medio alto, lo cual puede ser explicado por la presencia de la actividad minera. Por otro lado, la figura 6 sugiere que la mayor proporción de distritos que tienen PIB por km² alto se concentra en los departamentos que tienen costa. Esta figura también sugiere que los distritos que generan altos niveles de PIB por km² se encuentran distribuidos a lo largo de todo el territorio peruano. Estos resultados sugieren que el PIB por km² podría ser un indicador clave para medir el desempeño productivo a nivel regional, e incluso puede ser mejor indicador que el PIB pc.

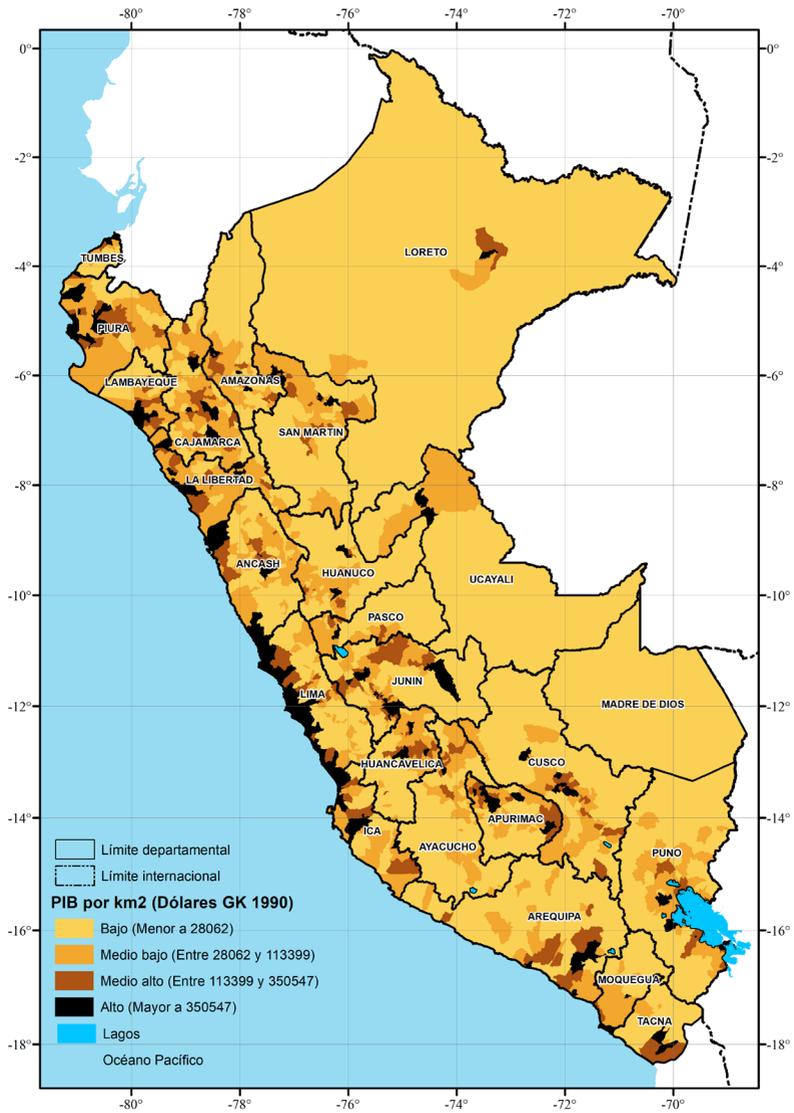
Las figuras descritas previamente sugieren que hay diferencias importantes en la distribución espacial del PIB pc y PIB por km². Por ello, es importante considerar ambos indicadores para analizar e identificar posibles patrones de desarrollo regional en el Perú. El PIB pc podemos considerarlo como un *proxy* del nivel de riqueza de las regiones, mientras que el PIB por km², como un *proxy* de la capacidad productiva de las regiones.

Figura 5
 PIB pc a nivel de distritos, 2018



Fuente: elaboración propia.

Figura 6
PIB por km² a nivel de distritos, 2018



Fuente: elaboración propia.

4.2 Desigualdad regional

Utilizando los datos a nivel distrital y el índice de Gini, analizamos la desigualdad regional del PIB pc (PIB/población) y del PIB por km² (PIB/área) con y sin Lima Metropolitana (LM) en el período 1993-2018⁶. Por un lado, el PIB pc busca medir el nivel de riqueza de cada región. Esto implica que el Gini del PIB pc es un indicador de concentración de la riqueza (desigualdad de la riqueza). Por otro lado, el PIB por km² busca medir la capacidad productiva del territorio de cada región. Es decir, el Gini del PIB por km² busca medir la concentración de la producción (desigualdad de la capacidad productiva).

Los resultados del Gini del PIB pc sugieren que la región de LM ha desempeñado un rol importante en la evolución de la desigualdad regional entre distritos. Por un lado, cuando incluimos a los distritos de LM, el Gini pasa de 0,3765 en 1993 a 0,3877 en 2018, es decir, se ha incrementado en un 2,96%. Por otro lado, cuando excluimos a los distritos de LM, el Gini pasa de 0,3048 en 1993 a 0,3020 en 2018, es decir, ha disminuido en un 0,9%. En la figura 7, mostramos la evolución del Gini del PIB pc entre distritos para el período 1993-2018. De acuerdo con esta figura, el Gini del PIB pc entre distritos con y sin LM no parece seguir el mismo patrón. Por ejemplo, entre 1993 y 1997, el Gini con LM venía aumentando, mientras que el Gini sin LM venía disminuyendo. Sin embargo, después del Fenómeno de El Niño (1997-1998) y la crisis asiática (1997-1999), el Gini con y sin LM disminuyó. No obstante, a partir de 1999, el Gini sin LM aumentó en mayor proporción que el Gini con LM hasta el año 2002; a partir de esta fecha, el Gini sin LM parece haber seguido la dinámica del Gini nacional. En esta figura también podemos notar que la crisis financiera internacional de 2008 disminuyó el Gini con y sin LM. Estos resultados sugieren que los períodos de crisis internacionales disminuyeron la desigualdad regional del PIB pc en todas las regiones del país. Asimismo, los resultados indican que el Gini con LM alcanzó su valor máximo de 0,4031 en 2011, mientras que el Gini sin LM alcanzó su valor máximo de 0,3199 en 2008. Después de alcanzar su valor máximo, el Gini con y sin LM muestra una tendencia negativa.

De otro lado, el Gini del PIB por km² con LM ha pasado de 0,9240 en 1993 a 0,9563 en 2018, es decir, se ha incrementado en un 3,48%; mientras que el Gini sin LM ha pasado de 0,8370 en 1993 a 0,8901 en 2018, es decir, se ha incrementado en 6,34%. Estos resultados sugieren que el Gini del PIB

⁶ Lima Metropolitana es un área metropolitana peruana conformada por distritos de la Provincia de Lima y la Provincia Constitucional del Callao.

por km^2 se ha incrementado tanto en LM como en el resto del país. En la figura 8, mostramos la evolución del Gini del PIB por km^2 con y sin LM. De acuerdo con esta figura, el Gini del PIB por km^2 ha crecido sostenidamente y no parece alterarse en los períodos de crisis económicas o financieras. Estos resultados sugieren que la concentración de la actividad económica medida por el Gini del PIB por km^2 ha aumentado sostenidamente tanto en LM como en el resto del país.

En concreto, los resultados sugieren que, entre 1993 y 2018, la desigualdad regional del ingreso medida por el Gini del PIB pc está descendiendo tanto en LM como en el resto del país, mientras que la concentración de la actividad económica medida por el Gini del PIB por km^2 está aumentando en todo el país. El aumento de la concentración de la actividad económica podría estar asociado con el aumento de la productividad y el crecimiento solo en algunos distritos del Perú. Esta situación podría estar configurando un sistema de centros y periferias que no se están articulando adecuadamente debido a factores geográficos y diseños inadecuados de políticas para el desarrollo regional. Por ello, futuras investigaciones podrían analizar la dinámica de la concentración de la actividad económica y de la población en regiones con similares características geográficas e históricas (Seminario & Palomino, 2021).

Figura 7
Evolución de la desigualdad regional del PIB pc, 1993-2018

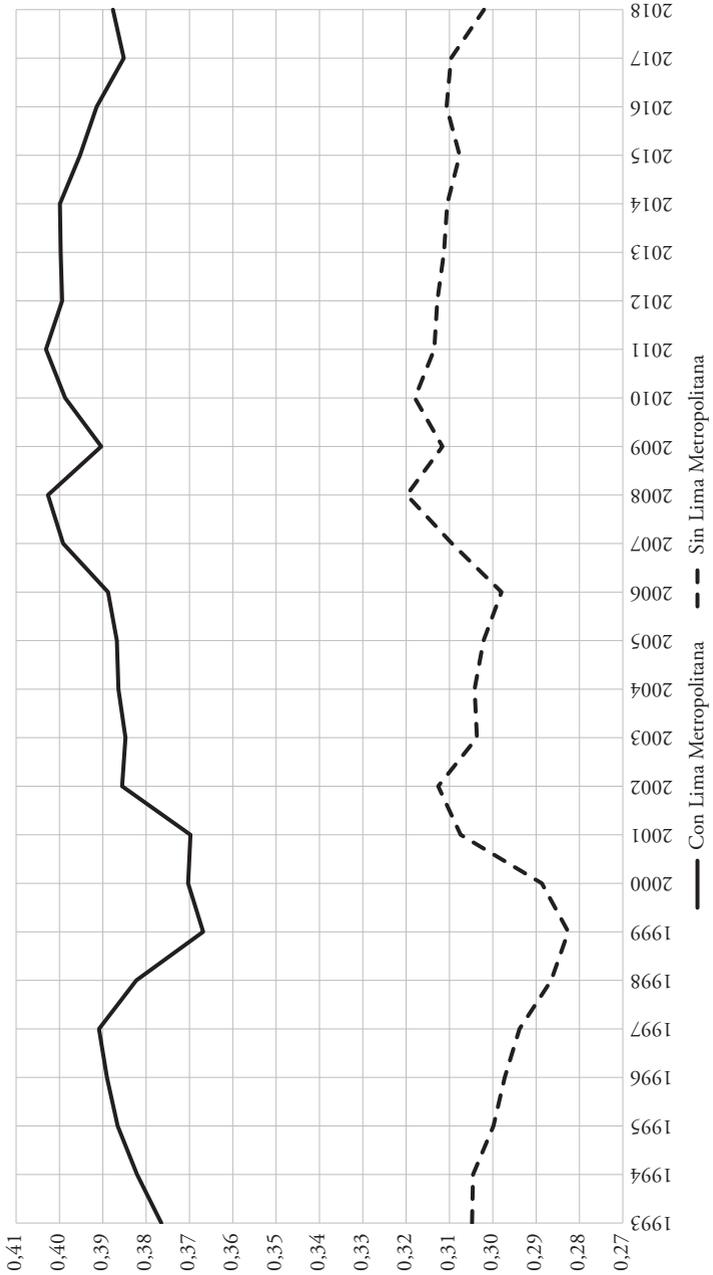
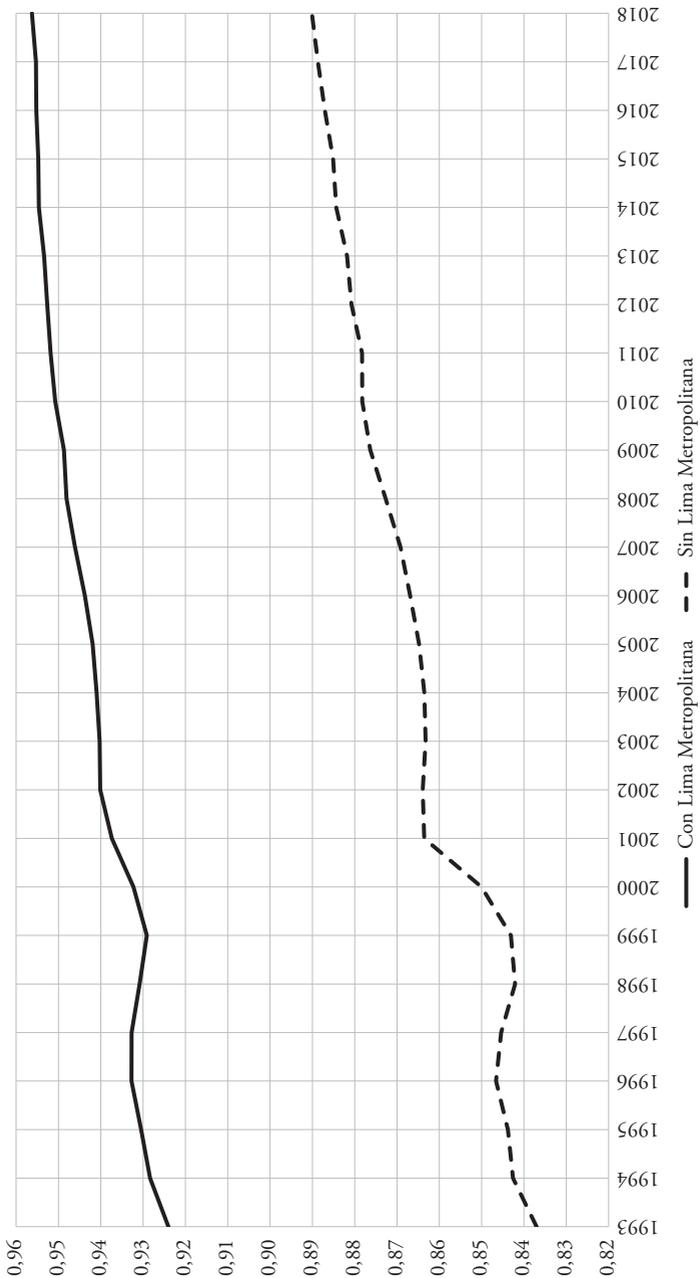


Figura 8
Evolución de la desigualdad regional del PIB por km², 1993-2018



4.3 Desarrollo regional

De acuerdo con Simon Kuznets, en las etapas iniciales del desarrollo económico de una nación existirán bajos niveles de desigualdad, en las fases intermedias existirán niveles de desigualdad altos y en las fases finales existirá por fin una mayor equidad. Esta relación es conocida como la curva en forma de U invertida de Kuznets.

En la figura 9, mostramos la relación entre el PIB pc real nacional y la desigualdad regional medida por el Gini del PIB pc real distrital. Esta figura sugiere que, entre 1993 y 2010, los distritos del Perú han incrementado su nivel de desarrollo, pero lo han hecho aumentando la desigualdad regional; mientras que, a partir de 2010, los distritos del Perú se están desarrollando disminuyendo la desigualdad regional.

Los resultados de la figura 9 sugieren la existencia de una curva de Kuznets que coincide con el ciclo de los precios de los metales que tiende a concentrar el PIB nominal en pocos distritos, pero nosotros aislamos el efecto de los precios de los metales porque utilizamos el PIB real en vez del PIB nominal. Sin embargo, consideramos que es necesario realizar un análisis de robustez para indicar si se valida o no la hipótesis de Kuznets en un período de 26 años. No obstante, el uso del PIB pc real como indicador único de desarrollo puede dar lugar a conclusiones erróneas, ya que este indicador no nos indica de qué manera el ingreso se distribuye entre los ciudadanos de un país y, por lo tanto, no permite medir el nivel de bienestar de una sociedad. Una alternativa para medir el bienestar de las distintas regiones de un país es el IDH, el cual combina indicadores de salud, educación y riqueza (medida por el PIB pc de paridad). Sin embargo, el problema es que el IDH no se calcula todos los años, ya que no se dispone de información censal todos los años. Por ello, proponemos un indicador alternativo al IDH para medir el desarrollo regional, el cual combina el PIB pc real como *proxy* de riqueza en línea con el IDH, densidad poblacional (población/área) como *proxy* de disponibilidad de fuerza laboral, y nivel de urbanización (población urbana / población total) como *proxy* del nivel de acceso a los servicios de salud-educación.

Utilizando la metodología propuesta por Ivanović (1963) y los datos a nivel de distritos, computamos un indicador alternativo para medir el desarrollo regional. En la figura 10, mostramos el nivel de desarrollo de todos los distritos para el año 2017. En esta figura, el color más oscuro indica mayor nivel de desarrollo, mientras que el color más claro indica menor nivel de desarrollo. De acuerdo con esta figura, los mayores niveles de desarrollo se concentran en los departamentos que tienen costa. Los resultados de esta figura

también sugieren que los distritos que son capitales provinciales son aquellos que tienen mayores niveles de desarrollo respecto a los otros distritos en los departamentos que no tienen costa. Un resultado al que vale la pena prestar atención son los niveles de desarrollo de algunos distritos en los departamentos con presencia de actividad minera. Por ejemplo, en esta figura, observamos que el departamento de Apurímac tiene distritos con altos niveles de desarrollo. Sin embargo, las estadísticas del PNUD (2019) también sugieren que algunos distritos del departamento de Apurímac han incrementado su nivel de IDH. Por lo tanto, nuestros resultados parecen ser consistentes con el IDH calculado por el PNUD a nivel distrital, pero, a diferencia del IDH, que solo se calcula para algunos años, nuestro indicador de desarrollo se puede calcular para todo el período de análisis.

En concreto, encontramos que los niveles de desarrollo a nivel de distritos se concentran en la Costa, en los distritos que son capitales provinciales y algunos distritos con actividad minera principalmente en el centro y sur del país. Asimismo, nuestros resultados parecen ser consistentes con el IDH calculado por el PNUD. Por lo tanto, futuras investigaciones podrían computar niveles de desarrollo regional en otros países utilizando la metodología propuesta en esta investigación. Sin embargo, para verificar la estabilidad del indicador de desarrollo, sobre todo en los distritos mineros, también podrían computar indicadores de desarrollo utilizando variantes del método que hemos utilizado (Pena-Trapero, 2009).

Figura 9
 Curva de Kuznets: PIB per cápita vs. Gini del PIB pc a nivel de distritos

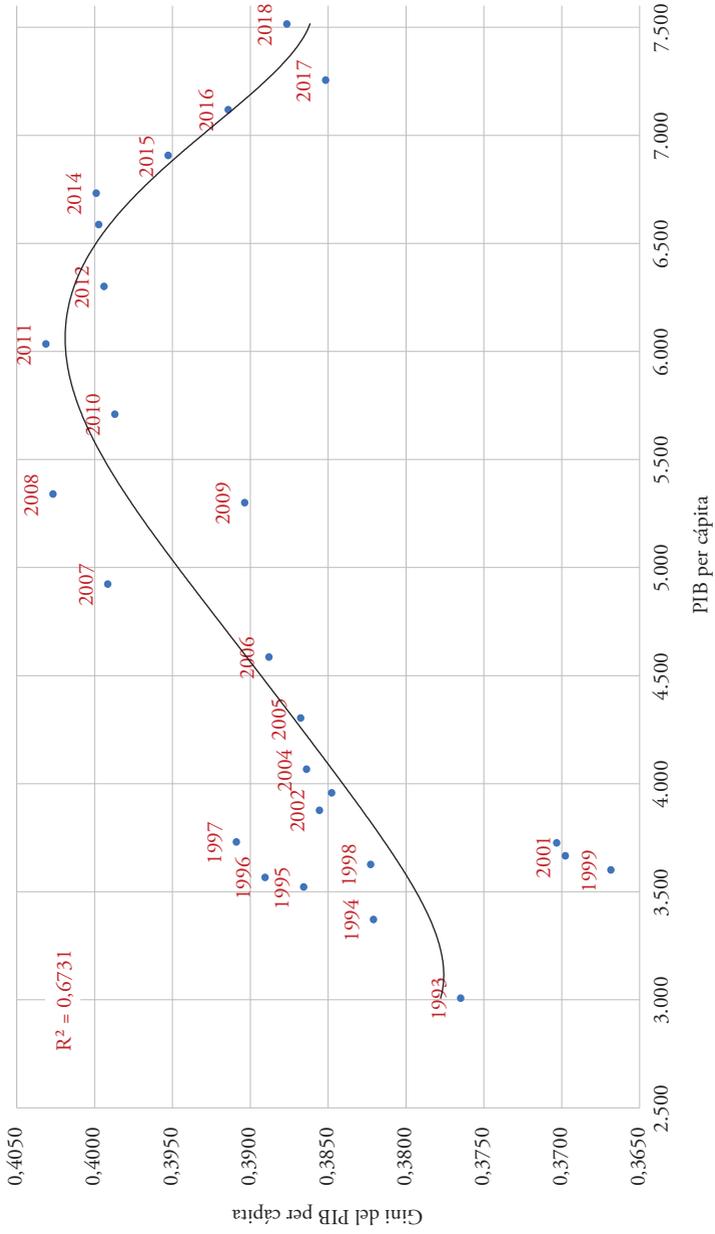
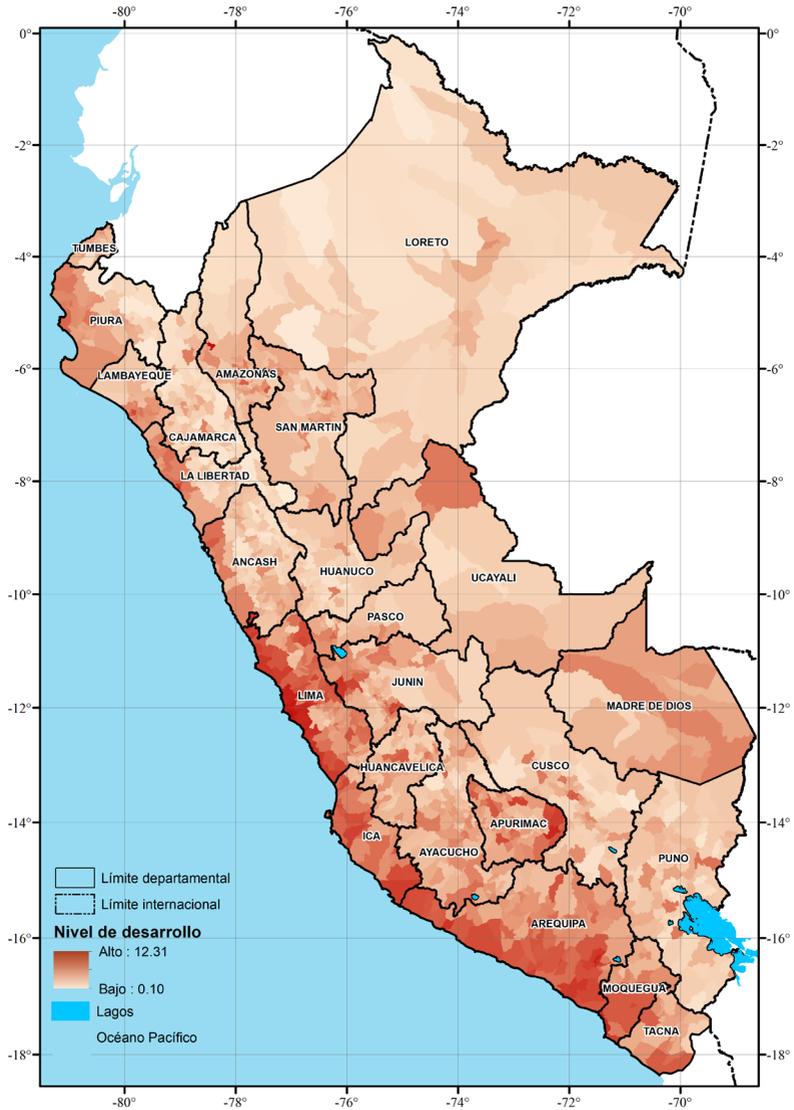


Figura 10
Nivel de desarrollo a nivel de distritos, 2017



Fuente: elaboración propia.

5. Conclusiones

El PIB es la variable más importante en los análisis de crecimiento económico. Sin embargo, a menudo se cometen errores de medición, en especial en los países en vías de desarrollo (Henderson *et al.*, 2012). En este contexto, las imágenes satelitales de las luces nocturnas proporcionan un medio alternativo para medir la actividad económica (Ghosh *et al.*, 2010). Sin embargo, los estudios que han usado las imágenes satelitales de luminosidad para medir la actividad económica en los países de Latinoamérica son escasos (Guerrero & Mendoza, 2019; Bickenbach *et al.*, 2016), a pesar de que los países de la región no cuentan con datos de calidad de PIB a nivel de subdivisiones administrativas territoriales. Por ello, el objetivo de esta investigación es ofrecer una metodología alternativa para estimar el PIB a nivel subnacional combinando información de censos y encuestas, estadísticas de cuentas nacionales y datos satelitales de luminosidad. Consideramos que el uso de esta metodología es relevante para países que no cuentan con información de calidad sobre la actividad económica a nivel subregional.

La metodología desarrollada para generar el PIB de divisiones territoriales que no cuentan con estadísticas de cuentas nacionales requiere información de cuatro variables: (1) densidad poblacional, (2) ingresos de los hogares, (3) series de PIB por departamento y (4) datos satelitales de luminosidad. En primer lugar, los datos de densidad poblacional los hemos generado utilizando información de los censos de población. En segundo lugar, la información de los ingresos de los hogares la hemos obtenido del PNUD. En tercer lugar, los datos del PIB por departamento los hemos obtenido del INEI y de Seminario *et al.* (2019). En cuarto lugar, los datos satelitales de luminosidad los hemos obtenido de NGDC-NOAA.

La estrategia para estimar el PIB de regiones administrativas pequeñas que no cuentan con estadísticas de cuentas nacionales consiste en estimar el PIB de

regiones administrativas grandes y luego usar estos datos para estimar el PIB de regiones administrativas pequeñas. En el Perú, contamos con estadísticas de actividad económica a nivel departamental, pero no a nivel provincial o distrital; por ello, para estimar el PIB a nivel distrital, primero debemos estimar el PIB a nivel provincial ajustado al PIB de los departamentos.

La estimación del PIB de regiones territoriales pequeñas consta de cuatro etapas: (1) generar series de luminosidad y densidad poblacional, (2) estimar el PIB de un año base, (3) estimar los parámetros de una función de producción, (4) generar series de PIB para todo el período de análisis, (5) ajustar las series de PIB estimados de las regiones territoriales menores al PIB de las regiones territoriales mayores.

En primer lugar, para generar series de luminosidad, usamos imágenes satelitales, y para generar series de densidad poblacional usamos los censos nacionales de población. Por un lado, las imágenes satelitales se diferencian en dos períodos: (1) 1993-2013 y (2) 2013-2018. Los datos del primer período tienen frecuencia anual y tienen una resolución de 30 segundos de arco (aproximadamente 1 km cuadrado), mientras que los datos del segundo período tienen frecuencia mensual y tienen una resolución de 15 segundos de arco (aproximadamente 0,000779 kilómetros cuadrados). Por otro lado, para generar las series de densidad poblacional utilizamos los censos de 1993, 2007 y 2017.

En segundo lugar, generamos series de PIB a nivel de provincias/distritos para el año base 2007 utilizando el método adaptado de Geary y Stark y la información de ingresos de los hogares derivada por el PNUD. De no contar con estadísticas de ingresos a nivel de subdivisiones territoriales administrativas, es necesario reproducir la metodología de Elbers *et al.* (2003) o del PNUD para generar estadísticas de ingresos con representatividad a nivel de las subdivisiones territoriales que queremos analizar.

En tercer lugar, estimamos una función de producción en la que el nivel de luminosidad es un *proxy* de capital y la densidad poblacional es un *proxy* de fuerza laboral. Los resultados sugieren que la función de producción adecuada para estimar el PIB es una función cúbica logarítmica donde la densidad poblacional es de grado uno y el nivel de luminosidad es de grado uno, dos y tres. Asimismo, los resultados sugieren que el nivel de luminosidad y la densidad poblacional explican mejor el comportamiento del PIB por km². Por ello, usamos los parámetros del modelo empírico que considera la interacción de la densidad poblacional y el nivel de luminosidad en una función cúbica y el PIB por km² como la variable dependiente.

En cuarto lugar, utilizando los coeficientes estimados en el paso anterior, estimamos el PIB a nivel de provincias y distritos para el período 1993-2018. Por último, en quinto lugar, utilizando la versión modificada del método de Geary y Stark, ajustamos los datos del PIB de regiones territoriales menores con datos del PIB de regiones territoriales mayores. Por un lado, usamos las series del PIB a nivel de departamentos obtenidas del INEI para ajustar las series estimadas del PIB a nivel de provincias. Por otro lado, usamos las series finales del PIB de las provincias para ajustar las series del PIB estimadas a nivel de distritos.

Utilizando los datos generados y mapas georreferenciados de la distribución espacial del PIB, encontramos que gran parte de los distritos que han generado mayor nivel de PIB se localizan en la Costa, algunos en la Sierra y muy pocos en la Selva. En la Costa, el departamento de Lima es aquel que concentra la mayor cantidad de distritos que generan altos niveles de PIB. En la Sierra, los distritos que generan mayor nivel de PIB se encuentran principalmente en las regiones mineras. En la Selva, gran parte de los distritos que generan mayor nivel de PIB se localizan en la Yunga Fluvial y Selva Alta, mientras que en la Selva Baja los distritos que generan altos niveles de PIB se localizan en las capitales departamentales.

Asimismo, utilizando las estadísticas del PIB, analizamos las tendencias recientes de desarrollo regional. Por un lado, adaptando la clasificación de países del Banco Mundial, encontramos que la mayor proporción de distritos pertenecen al grupo de PIB pc medio bajo y PIB por km² bajo. Los resultados también sugieren que hay diferencias importantes en la distribución espacial del PIB pc y PIB por km². Por ello, es importante considerar ambos indicadores para analizar e identificar posibles patrones de desarrollo regional en el Perú.

Por otro lado, utilizando los datos a nivel distrital, encontramos que la desigualdad regional del ingreso medido por el Gini del PIB pc está descendiendo tanto en LM como en el resto del país, mientras que la concentración de la actividad económica medida por el Gini del PIB por km² está aumentando en todo el país. El aumento en la concentración de la actividad económica podría estar asociado con el aumento de la productividad y el crecimiento solo en algunos distritos del Perú. Esta situación podría estar configurando un sistema de centros y periferias que no se están articulando adecuadamente debido a factores geográficos y diseños inadecuados de políticas para el desarrollo regional.

De otro lado, los datos sugieren que, entre 1993 y 2010, los distritos del Perú han incrementado su nivel de desarrollo, pero lo han hecho aumentando

la desigualdad regional; mientras que, a partir de 2010, los distritos del Perú se están desarrollando disminuyendo la desigualdad regional. Sin embargo, consideramos que es necesario realizar un análisis de robustez para indicar si se valida o no la hipótesis de Kuznets en un período de 26 años. Adicionalmente, dado que el PIB pc no es suficiente para medir el bienestar de un país, utilizamos los datos a nivel de distritos y la metodología propuesta por Ivanović (1963) para computar un indicador alternativo de desarrollo regional. Utilizando este nuevo indicador, encontramos que los altos niveles de desarrollo a nivel de distritos se concentran en la Costa, en los distritos que son capitales provinciales y algunos distritos con actividad minera principalmente en el centro y sur del país. Asimismo, encontramos que nuestro indicador de desarrollo regional parece ser consistente con el IDH calculado por el PNUD a nivel de distritos.

6. Limitaciones y futuras líneas de investigación

En esta investigación, estimamos una función de producción cúbica en la que el nivel de luminosidad es *proxy* del capital y la densidad poblacional es *proxy* de la fuerza laboral. Utilizamos esta función de producción porque encontramos que esta forma funcional reproduce bastante bien las series de PIB a nivel departamental en el Perú. Sin embargo, futuras investigaciones podrían analizar la estabilidad de esta forma funcional realizando estimaciones en otros países y utilizando diferentes divisiones territoriales. Las futuras investigaciones también podrían realizar estimaciones utilizando otras variables en la función de producción. Por ejemplo, podrían incorporar variables geográficas como el uso del suelo, la pendiente del territorio, el caudal de los ríos, el nivel de precipitación, el nivel de temperatura, etc.

Los parámetros estimados en la función de producción a nivel de provincias y distritos consideran la información del año 2007. Elegimos este año porque es el principal año base de las estadísticas económicas del Perú. Sin embargo, futuras investigaciones podrían estimar la función de producción utilizando información de diferentes años censales que no estaban disponibles a nivel de provincias y distritos en el momento de iniciar esta investigación. Asimismo, futuras investigaciones podrían utilizar la metodología desarrollada en esta investigación para estimar el PIB de divisiones territoriales pequeñas de otros países que no cuentan con estadísticas de cuentas nacionales. Por ejemplo, Zegarra *et al.* (2020, 2021) utilizaron esta metodología para estimar el PIB de cada una de las islas en Las Bahamas y con ello analizaron los impactos de los huracanas en dicho lugar.

Utilizando los datos a nivel de distritos e información de población de los CP de 2007, generamos series de PIB a nivel de CP y *grids*. Sin embargo, futuras investigaciones podrían usar la información de varios censos de población a nivel de CP para estimar el PIB a nivel de *grids*. Asimismo, futuras investi-

gaciones podrían utilizar los píxeles o *grids* de luminosidad para distribuir el PIB distrital en todo el período de análisis. La información a nivel de *grids* podría permitir otro tipo de ejercicios. Por ejemplo, Seminario *et al.* (2019), utilizando información a nivel de *grids*, generaron indicadores de actividad económica de las regiones naturales y cuencas hidrográficas desde 1795.

Las investigaciones futuras podrían considerar mejorar la calidad de las imágenes satelitales de noche. El uso de diodos emisores de luz (LED por sus siglas en inglés) hace que las imágenes de VIIRS puedan parecer más oscuras de lo que realmente son (De Miguel *et al.*, 2021; Kyba *et al.*, 2017). La diferente hora de paso del DMSP y VIIRS también podría estar generando problemas en la calidad de las imágenes satelitales de luminosidad. Por lo tanto, es necesario tratar la calidad de las imágenes satelitales tomadas durante la noche de acuerdo con el contexto y tipo de economía (Wang *et al.*, 2021).

Futuras investigaciones también podrían explorar otras formas para completar los MV de las regiones que no tienen información del nivel de luminosidad debido a la falta de información satelital. Nosotros asumimos que el nivel de luminosidad de una región se comporta como el promedio de sus regiones colindantes. Sin embargo, es posible que ciertas regiones colindantes no tengan conexión debido a la geografía y el relieve. Por lo tanto, nuestra forma de completar los MV podría estar generando sesgos hacia las regiones con mayor nivel de luminosidad pero que no tienen conexión con la región de interés que no cuenta con información de luminosidad.

Las futuras investigaciones podrían estimar funciones de producción que combinen la información de las imágenes satelitales tomadas en la noche y las imágenes satelitales tomadas en el día (Huang, Hsiang, & González-Navarro, 2021; Engstrom, Hersh, & Newhouse, 2017; Jean *et al.*, 2016). La combinación de la información de las imágenes satelitales de día y de noche será relevante porque diversos estudios han encontrado que las luces de noche son fuente de contaminación (Levin *et al.*, 2020) y pueden estar amenazando la biodiversidad de las especies (Pauwels *et al.*, 2019). En este contexto, se están dando diversas iniciativas para usar luces más ecológicas; por ejemplo, se viene recomendando el uso de LED (Pawson & Bader, 2014). En concreto, mientras sigamos utilizando los mismos sistemas para iluminarnos en la noche, las imágenes satelitales tomadas durante la noche serán útiles para estimar el PIB. Sin embargo, cuando los sistemas de iluminación alrededor del mundo cambien, será necesario recurrir a nuevos métodos y a nueva información para seguir generando estadísticas de actividad económica para las regiones que no cuentan con información de cuentas nacionales.

Finalmente, futuras investigaciones también podrían combinar las imágenes satelitales con información georreferenciada de teléfonos celulares para estimar PIB sectorial a nivel regional. Por un lado, utilizando la densidad poblacional y asignando información del nivel de luminosidad a las regiones sin información satelital, hemos podido generar series de PIB consistentes con la historia en las regiones agrícolas o ganaderas. Sin embargo, es importante mencionar que las distintas actividades económicas a nivel regional tienen distintos niveles de uso de luz eléctrica reflejado en las imágenes satelitales tomadas durante la noche. Por ello, el uso adecuado de las imágenes satelitales tomadas durante el día podría mejorar nuestras estimaciones de PIB en las regiones agrícolas o ganaderas. Por otro lado, las imágenes satelitales tomadas durante el día y la información georreferenciada de teléfonos celulares podrían permitir generar indicadores de actividad económica sectorial a nivel regional y de alta frecuencia (mensual, semanal o diaria). Sin embargo, es necesario mencionar que uso de los datos georreferenciados de teléfonos celulares debe resguardar la privacidad de los usuarios (Aromi *et al.*, 2020, 2021; Couture *et al.*, 2021; Cook, Currier, & Glaser, 2022).

Referencias

- Amavilah, V. H. (2018). Artificial nighttime lights and the «real» well-being of nations: Measuring economic growth from outer space and welfare from right here on Earth. *Journal of Economics and Political Economy*, 5(2), 209-218.
- Aromi, D., Bonel, M. P., Cristia, J. P., Llada, M., & Palomino, L. (2021). *Socioeconomic status and mobility during the COVID-19 pandemic: An analysis of eight large Latin American cities*. Inter-American Development Bank, Working Paper 1184.
- Aromi, D., Bonel, M. P., Cristia, J., Llada, M., Pereira, J. I., Pulido, X., & Santamaría, J. (2020). *#QuédateEnCasa: Las políticas de distanciamiento social y la movilidad en América Latina y el Caribe*. Inter-American Development Bank, Working Paper 1147.
- Badia-Miró, M. (2015). The evolution of the location of economic activity in Chile in the long run: A paradox of extreme concentration in absence of agglomeration economies. *Estudios de Economía*, 42(2).
- Bennie, J., Davies, T. W., Duffy, J. P., Inger, R., & Gaston, K. J. (2014). Contrasting trends in light pollution across Europe based on satellite observed nighttime lights. *Scientific Reports*, 4(1), 1-6.
- Bennie, J., Davies, T. W., Inger, R., & Gaston, K. J. (2014). Mapping artificial lightscares for ecological studies. *Methods in Ecology and Evolution*, 5(6), 534-540.
- Bickenbach, F., Bode, E., Nunnenkamp, P., & Söder, M. (2016). Night lights and regional GDP. *Review of World Economics*, 152(2), 425-447.
- 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.
- Clark, H., Pinkovskiy, M., & Sala-i-Martin, X. (2017). *China's GDP growth may be understated* (N.º w23323). National Bureau of Economic Research.
- Comisión Económica para América Latina y el Caribe (Cepal). (2015). *Panorama del desarrollo territorial en América Latina y el Caribe, 2015: pactos para la igualdad territorial*.
- Cook, C., Currier, L., & Glaeser, E. L. (2022). *Urban mobility and the experienced isolation of students and adults* (N.º w29645). National Bureau of Economic Research. <https://www.nber.org/papers/w29645>

- Couture, V., Dingel, J. I., Green, A., Handbury, J., & Williams, K. R. (2021). JUE Insight: Measuring movement and social contact with smartphone data: A real-time application to COVID-19. *Journal of Urban Economics*, 103328. <https://doi.org/10.1016/j.jue.2021.103328>
- Crafts, N. (2004). *Regional GDP in Britain, 1871-1911: Some estimates*. Department of Economic History, London School of Economics and Political Science, Londres, Reino Unido.
- De Miguel, A. S., Zamorano, J., Aubé, M., Bennie, J., Gallego, J., Ocaña, F., ..., & Gaston, K. J. (2021). Colour remote sensing of the impact of artificial light at night (II): Calibration of DSLR-based images from the International Space Station. *Remote Sensing of Environment*, 264, 112611.
- Doll, C. N., Muller, J. P., & Morley, J. G. (2006). Mapping regional economic activity from night-time light satellite imagery. *Ecological Economics*, 57(1), 75-92.
- Elbers, C., Lanjouw, J. O., & Lanjouw, P. (2003). Micro-level estimation of poverty and inequality. *Econometrica*, 71(1), 355-364.
- Elvidge, C. D., Baugh, K. E., Anderson, S. J., Sutton, P. C., & Ghosh, T. (2012). The Night Light Development Index (NLDI): A spatially explicit measure of human development from satellite data. *Social Geography*, 7(1), 23-35.
- Elvidge, C. D., Sutton, P. C., Ghosh, T., Tuttle, B. T., Baugh, K. E., Bhaduri, B., & Bright, E. (2009). A global poverty map derived from satellite data. *Computers & Geosciences*, 35(8), 1652-1660.
- Engstrom, R., Hersh, J., & Newhouse, D. (2017). *Poverty from space: Using high-resolution satellite imagery for estimating economic well-being*. Policy Research Working Paper 8284. World Bank.
- Geary, F., & Stark, T. (2002). Examining Ireland's post-famine economic growth performance. *The Economic Journal*, 112(482), 919-935.
- Ghosh, T. L., Powell, R. D., Elvidge, C. E., Baugh, K. C., Sutton, P., & Anderson, S. (2010). Shedding light on the global distribution of economic activity. *The Open Geography Journal*, 3(1).
- Guerrero, V. M., & Mendoza, J. A. (2019). On measuring economic growth from outer space: A single country approach. *Empirical Economics*, 57(3), 971-990.
- Henderson, J. V., Squires, T., Storeygard, A., & Weil, D. (2018). The global distribution of economic activity: Nature, history, and the role of trade. *The Quarterly Journal of Economics*, 133(1), 357-406.
- Henderson, J. V., Storeygard, A., & Weil, D. N. (2012). Measuring economic growth from outer space. *American Economic Review*, 102(2), 994-1028.
- Henderson, V., Storeygard, A., & Weil, D. N. (2011). A bright idea for measuring economic growth. *American Economic Review*, 101(3), 194-99.
- Huang, L. Y., Hsiang, S., & González-Navarro, M. (2021). *Using satellite imagery and deep learning to evaluate the impact of anti-poverty programs*. <https://arxiv.org/abs/2104.11772>

- Instituto Nacional de Estadística e Informática (INEI). (2018). *Perú: perfil sociodemográfico. Informe nacional. Censos Nacionales 2017: XII de Población, VII de Vivienda y III de Comunidades Indígenas*.
- Ivanović, B. (1963). Classification of underdeveloped areas according to level of economic development. *Eastern European Economics*, 2(1-2), 46-61.
- Jean, N., Burke, M., Xie, M., Davis, W. M., Lobell, D. B., & Ermon, S. (2016). Combining satellite imagery and machine learning to predict poverty. *Science*, 353(6301), 790-794.
- Kyba, C. C., Kuester, T., De Miguel, A. S., Baugh, K., Jechow, A., Hölker, F., ..., & Guanter, L. (2017). Artificially lit surface of Earth at night increasing in radiance and extent. *Science Advances*, 3(11), e1701528.
- Levin, N., & Zhang, Q. (2017). A global analysis of factors controlling VIIRS nighttime light levels from densely populated areas. *Remote Sensing of Environment*, 190, 366-382.
- Levin, N., Kyba, C. C., Zhang, Q., De Miguel, A. S., Román, M. O., Li, X., ..., & Wang, Z. (2020). Remote sensing of night lights: A review and an outlook for the future. *Remote Sensing of Environment*, 237, 111443.
- Michalopoulos, S., & Papaioannou, E. (2014). National institutions and subnational development in Africa. *The Quarterly Journal of Economics*, 129(1), 151-213.
- Michalopoulos, S., & Papaioannou, E. (2018). Spatial patterns of development: A meso approach. *Annual Review of Economics*, 10(1), 383-410.
- Nordhaus, W., & Chen, X. (2015). A sharper image? Estimates of the precision of nighttime lights as a proxy for economic statistics. *Journal of Economic Geography*, 15(1), 217-246.
- Otsu, N. (1979). A threshold selection method from gray-level histograms. *IEEE Transactions on Systems, Man, and Cybernetics*, 9(1), 62-66.
- Pauwels, J., Le Viol, I., Azam, C., Valet, N., Julien, J. F., Bas, Y., ..., & Kerbiriou, C. (2019). Accounting for artificial light impact on bat activity for a biodiversity-friendly urban planning. *Landscape and Urban Planning*, 183, 12-25.
- Pawson, S. M., & Bader, M. F. (2014). LED lighting increases the ecological impact of light pollution irrespective of color temperature. *Ecological Applications*, 24(7), 1561-1568.
- Pena-Trapero, B. (2009). La medición del bienestar social: una revisión crítica. *Estudios de Economía Aplicada*, 27(2), 299-324.
- Pinkovskiy, M., & Sala-i-Martin, X. (2016). Lights, camera... income! Illuminating the national accounts-household surveys debate. *The Quarterly Journal of Economics*, 131(2), 579-631.
- Pinkovskiy, M., & Sala-i-Martin, X. (2018). *Shining a light on purchasing power parities* (N.º w24419). National Bureau of Economic Research.

- Programa de las Naciones Unidas para el Desarrollo (PNUD). (2019). *El reto de la igualdad: una lectura a las dinámicas territoriales en el Perú. Datos disponibles para el período 2003-2019*. Recuperado de <https://www.pe.undp.org/content/peru/es/home/library/poverty/el-reto-de-la-igualdad.html>
- Seminario, B., & Palomino, L. (2021). Concentración de la actividad económica y de la población en el Perú, 1795-2018. En A. Beltrán, C. A. Sanborn & G. Yamada (Eds.), *En búsqueda de un desarrollo integral: 20 ensayos en torno al Perú del Bicentenario* (pp. 89-131). Lima: Universidad del Pacífico.
- Seminario, B., Zegarra, M. A., & Palomino, L. (2019). *Estimación del PIB departamental y análisis de la desigualdad regional en el Perú, 1795-2017*. IDB Working Paper 1016.
- Theil, H. (1967). *Economics and information theory* (N.º 04; HB74. M3, T4).
- Wang, W., Cheng, H., & Zhang, L. (2012). Poverty assessment using DMSP/OLS night-time light satellite imagery at a provincial scale in China. *Advances in Space Research*, 49(8), 1253-1264.
- Wang, Z., Román, M. O., Kalb, V. L., Miller, S. D., Zhang, J., & Shrestha, R. M. (2021). Quantifying uncertainties in nighttime light retrievals from Suomi-NPP and NOAA-20 VIIRS Day/Night Band data. *Remote Sensing of Environment*, 263, 112557.
- Xu, H., Yang, H., Li, X., Jin, H., & Li, D. (2015). Multi-scale measurement of regional inequality in Mainland China during 2005-2010 using DMSP/OLS night light imagery and population density grid data. *Sustainability*, 7(10), 13469-13499.
- Yu, B., Shi, K., Hu, Y., Huang, C., Chen, Z., & Wu, J. (2015). Poverty evaluation using NPP-VIIRS nighttime light composite data at the county level in China. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 8(3), 1217-1229.
- Zegarra, M. A., Giles, L. Gartner, M., & Palomino, L. (2021). *The macro-economic effects of hurricanes in The Bahamas: A case study using satellite night light luminosity*. Inter-American Development Bank, Catalogs and Brochures.
- Zegarra, M. A., Schmid, J. P., Palomino, L., & Seminario, B. (2020). *Impact of hurricane Dorian in The Bahamas: A view from the sky*. IDB Technical Note 1857.
- Zhao, M., Zhou, Y., Li, X., Cao, W., He, C., Yu, B., ..., & Zhou, C. (2019). Applications of satellite remote sensing of nighttime light observations: Advances, challenges, and perspectives. *Remote Sensing*, 11(17), 1971.

Anexos

Anexo 1: Estimación del PIB a nivel de centros poblados (CP)

Diferentes estudios indican que las ciudades más pobladas son las que generan mayores niveles de PIB. Por ello, creemos que podemos estimar el PIB de los CP usando la información de población del censo de población y vivienda de 2007. Nosotros suponemos que cantidad de CP se mantiene y que lo que genera la dinámica en el tiempo es el tamaño del CP, es decir, la cantidad de personas por CP. La estimación del PIB a nivel de CP tiene tres etapas:

1. Estimamos el PIB del CP en el año base.

Usando la información de población a nivel de CP, hemos construido el PIB a nivel de CP para el año 2007 utilizando la siguiente fórmula:

$$y_{i,k,t_0}^{E,cp} = y_{k,t_0}^{r,d} * \frac{p_{i,k,t_0}^{cp}}{\sum_{j=1}^N p_{j,k,t_0}^{cp}}$$

donde $y_{i,k,t_0}^{E,cp}$ corresponde al PIB estimado del CP i del distrito k en el año base, $y_{k,t_0}^{r,d}$ representa el PIB real del distrito k al que pertenece el CP i en el año base, p_{i,k,t_0}^{cp} corresponde a la población del CP i del distrito k en el año base, $\sum_{j=1}^N p_{j,k,t_0}^{cp}$ corresponde a la población total de todos los CP que pertenecen al distrito k . En este caso, k toma valores de 1 a 1.834, i toma valores de 1 a 98.011, mientras que los valores de N dependen de la cantidad de CP que tenga cada distrito.

2. Estimamos el PIB de los CP para el período 1993-2018 con base en el PIB estimado del año base y las tasas de crecimiento del PIB de los distritos utilizando la siguiente fórmula:

$$y_{i,k,t}^{E,cp} = y_{i,k,t_0}^{E,cp} * \frac{y_{k,t}^{r,d}}{y_{k,t_0}^{r,d}}$$

donde $y_{i,k,t}^{E,cp}$ corresponde al PIB estimado del CP i del distrito k en el período t , $y_{i,k,t_0}^{E,cp}$ indica el PIB estimado del CP i del distrito k en el año base, $y_{k,t}^{r,d}$ y $y_{k,t_0}^{r,d}$ representan el PIB real del distrito k en los períodos t y t_0 respectivamente.

- Utilizando la metodología de Geary y Stark (2002), ajustamos los datos de PIB de los CP con los datos de PIB de los distritos con la siguiente fórmula:

$$y_{i,k,t}^{r,cp} = y_{k,t}^{r,d} * \frac{y_{i,k,t}^{E,cp}}{\sum_{j=1}^N y_{j,k,t}^{E,cp}}$$

donde $y_{i,k,t}^{r,cp}$ representa el PIB real estimado y corregido para el CP i del distrito k en el período t , $y_{k,t}^{r,d}$ indica el PIB real del distrito k en el período t , $y_{i,k,t}^{E,cp}$ corresponde al PIB estimado del CP i del distrito k en el período t , y $\sum_{j=1}^N y_{j,k,t}^{E,cp}$ corresponde al PIB total estimado de los CP que pertenecen al distrito k .

Anexo 2: Estimación del PIB a nivel de *grids*

Uno de los principales problemas cuando se trabaja con unidades administrativas pequeñas es el cambio en su cantidad debido a la creación de nuevas unidades administrativas. Por ejemplo, en el año 2007, se registraron 1.834 distritos, mientras que en 2013 la cantidad aumentó a 1.876. Para evitar el problema de creación de nuevas unidades administrativas, se optó por dividir al Perú en *grids* (cuadrados rectangulares), a fin de poder analizar cualquier región sin importar su unidad administrativa. Por ejemplo, utilizando *grids*, hemos calculado el PIB de las regiones naturales y cuencas hidrográficas del Perú. El criterio del tamaño del *grid* responde al nivel de aglomeración continua que se observa en los CP. Realizando diferentes pruebas, Seminario, Zegarra y Palomino (2019) encontraron que la unidad de aglomeración continua de los CP es la legua. La legua es una antigua unidad de longitud que expresa la distancia que una persona, a pie, o en cabalgadura, puede recorrer durante una hora; es decir, es una medida itineraria. Dado que una persona recorre normalmente a pie una gama de distancias, la legua se mantiene en esa gama, pero según el tipo de terreno predominante en cada región. Usando el criterio de la legua, dividimos al Perú en 1.067.325 *grids*, donde el tamaño de cada *grid* es 1/5 de legua marina. La estimación del PIB a nivel de *grid* tiene dos etapas:

1. Estimamos el PIB a nivel de *grids* utilizando la siguiente fórmula:

$$y_{i,t}^{E,g} = \sum_{j=1}^N y_{j,i,t}^{r,cp}$$

donde $y_{i,t}^{E,g}$ representa el PIB estimado del *grid* i en el período t , y $y_{j,i,t}^{r,cp}$ corresponde al PIB real del CP j que pertenece al *grid* i en el período t . En este caso, i toma valores de 1 a 1.067.325 porque el Perú se ha dividido en esta cantidad de *grids*, mientras que el valor de N depende de la cantidad de CP que tiene cada *grid*.

2. Usando la metodología de Geary y Stark (2002), ajustamos los datos del PIB de los *grids* con los datos del PIB nacional.

Este ejercicio es necesario porque puede que un CP se encuentre justo en la frontera de un *grid* y, con ello, se genere una doble contabilidad. Por ello, los datos del PIB real de los *grids* los ajustamos a los datos nacionales con la siguiente fórmula:

$$y_{i,t}^{r,g} = y_t^{Nac} * \frac{y_{i,t}^{E,g}}{\sum_{j=1}^M y_{j,t}^{E,g}}$$

donde $y_{i,t}^{r,g}$ representa el PIB real estimado y corregido para el *grid* i en el período t , y_t^{Nac} corresponde al PIB real nacional en el período t , $y_{i,t}^{E,g}$ corresponde al PIB real estimado del *grid* i en el período t , y $\sum_{j=1}^M y_{j,t}^{E,g}$ representa el PIB estimado total de todos los *grids*. En este caso, i toma valores de 1 a 1.067.325 porque el Perú se ha dividido en esta cantidad de *grids* de 1/5 de legua marina cada uno, y M indica la cantidad total de *grids*. El anexo 3 muestra el mapa de *grids* con y sin PIB para el período 1993-2018.

Anexo 3: Grids con y sin PIB, 1993-2018



Fuente: elaboración propia.

Anexo 4: Datos de actividad económica a nivel de provincias y distritos del Perú

Los datos de actividad económica generados en esta investigación a nivel de provincias y distritos del Perú para el período 1993-2018 se encuentran almacenados [aquí](#).