

No dejar a nadie atrás: aplicaciones prácticas de SAE en América Latina

División de Estadísticas - CEPAL

Andrés Gutiérrez

2022

Objetivos de Desarrollo Sostenible



No dejar a nadie atrás



Los indicadores de los Objetivos de Desarrollo Sostenible deben desagregarse, cuando corresponda, por ingreso, sexo, edad, raza, etnia, estatus migratorio, discapacidad y ubicación geográfica, u otras características, de conformidad con los Principios Fundamentales de las Estadísticas Oficiales.

Marco de indicadores globales para los Objetivos de Desarrollo Sostenible (A/RES/71/313).

Limitaciones de las encuestas de hogares y el uso de información auxiliar

Estimadores directos

- El tamaño de la muestra en cada dominio no está planificado de antemano.
- Cualquier estimación de indicadores relativos (medias y proporciones) deberá hacer uso de un estimador de tipo razón: *numerador aleatorio y denominador aleatorio*.

$$\hat{\theta}_d^{Dir} = \frac{\sum_{k \in s_d} w_k y_k}{\sum_{k \in s_d} w_k} \quad \widehat{AV}(\hat{\theta}_d^{Dir}) = \sum_{k \in s_d} \sum_{l \in s_d} \frac{\Delta_{kl}}{\pi_{kl}} \frac{e_k}{\pi_k} \frac{e_l}{\pi_l}$$

Cuando el tamaño de la muestra $n_d = \#(s_d)$ no es lo suficientemente grande, ninguno de los estimadores anteriores será preciso ni consistente.

BADEHOG - Banco de Datos de Encuestas de Hogares de la CEPAL

Este es un insumo fundamental para los indicadores sociales armonizados producidos por la CEPAL y apoya las investigaciones y estudios de algunas agencias de las Naciones Unidas.

- Desarrollo principal en los años 90.
- Recopila anualmente datos de encuestas de hogares de 18 países de América Latina.
- Encuestas de hogares que miden el ingreso y se utilizan para estimar la pobreza y la desigualdad distributiva
- Encuestas desde 1980 en adelante.
- Una encuesta por año (período más cercano al final del año).

Banco de datos censales del CELADE

La División de Población de la CEPAL (CELADE) ha desarrollado un repositorio integral de datos censales.

- Redatam es un software de procesamiento estadístico especializado en microdatos de censos de población y vivienda.
- Redatam permite el procesamiento, análisis y difusión de datos, pudiendo manejar grandes volúmenes de datos, en estructuras jerárquicas con gran velocidad de procesamiento.
- A través de esta plataforma, tenemos a nuestra disposición todos los censos de la región.

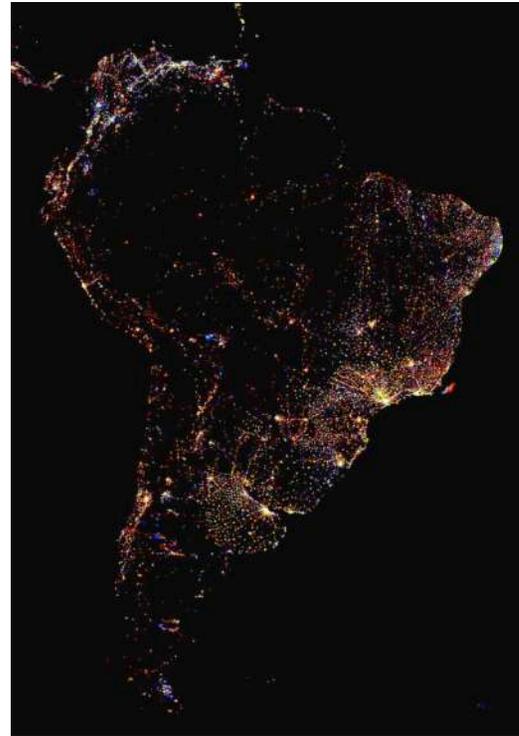
Imágenes de satélites

Accedemos a esta información a través de Google Earth Engine, que brinda facilidades para analizar y obtener estos datos a través de los lenguajes de programación Javascript y Python, y recientemente desde 2021 en R con el paquete *rgee*.

Entre las principales ventajas de la información basada en sensores remotos se encuentra la facilidad de acceso a información con alta cobertura geográfica que es imposible de obtener por medios tradicionales como encuestas o registros administrativos.

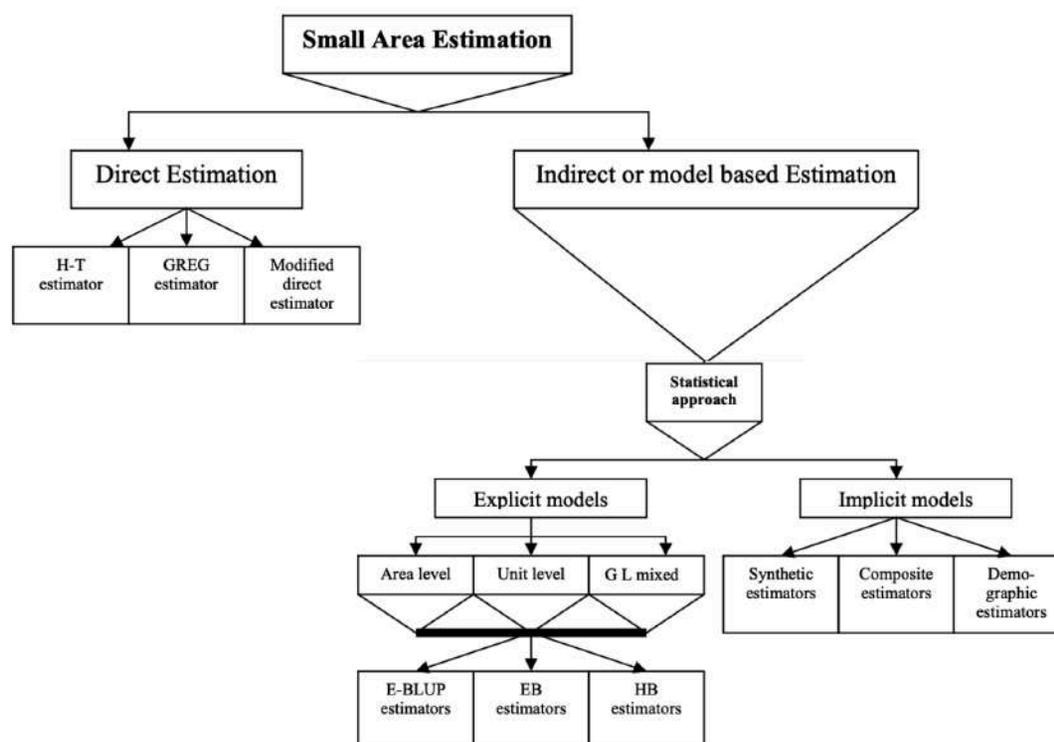
Imágenes de satélites

- Las luces nocturnas son un ejemplo notable del uso de imágenes de satélite y se ha utilizado como proxy para diferentes variables socioeconómicas.
- Las fracciones de cobertura urbana y de cultivos brindan a los países la opción de compensar la falta de censos de población y encuestas detalladas.



Algunos modelos SAE en CEPAL

Alternativas para hacer SAE



Fuente: adaptación de Rahman (2008).

Dos tipos de métodos y dos tipos de enfoques

Los estimadores SAE utilizados en la CEPAL se pueden dividir en dos tipos principales:

1. Estimadores basados en modelos de área.
2. Estimadores basados en modelos de unidad.

La elección del método se realiza en función del nivel en el que se encuentra la información auxiliar (a nivel de dominio o de agregación - a nivel de hogar o de

Hay dos enfoques inferenciales:

1. Frecuentista.
2. Bayesiana.

Ultimamente nos hemos inclinado por el enfoque bayesiano, por su facilidad de interpretación y su eficiencia a la hora de producir estimaciones del RMSE.

Metodología general para SAE

- Estandarización y homologación de covariables en las bases de datos (censos y encuestas de hogares).
- Actualización de conteos intercensales preservando las estructuras censales teniendo en cuenta las marginales actualizadas de la encuesta de hogares.
- Definición de los modelos para los indicadores de interés, considerando
- Predicción en postestratos censales y áreas pequeñas, junto con la estimación del MSE (réplicas Bootstrap o MCMC).
- Validación de supuestos del modelo y benchmarking utilizando estimaciones de CEPAL de ingreso medio y pobreza a nivel nacional, urbano y rural.
- Generación de mapas con estimaciones puntuales y

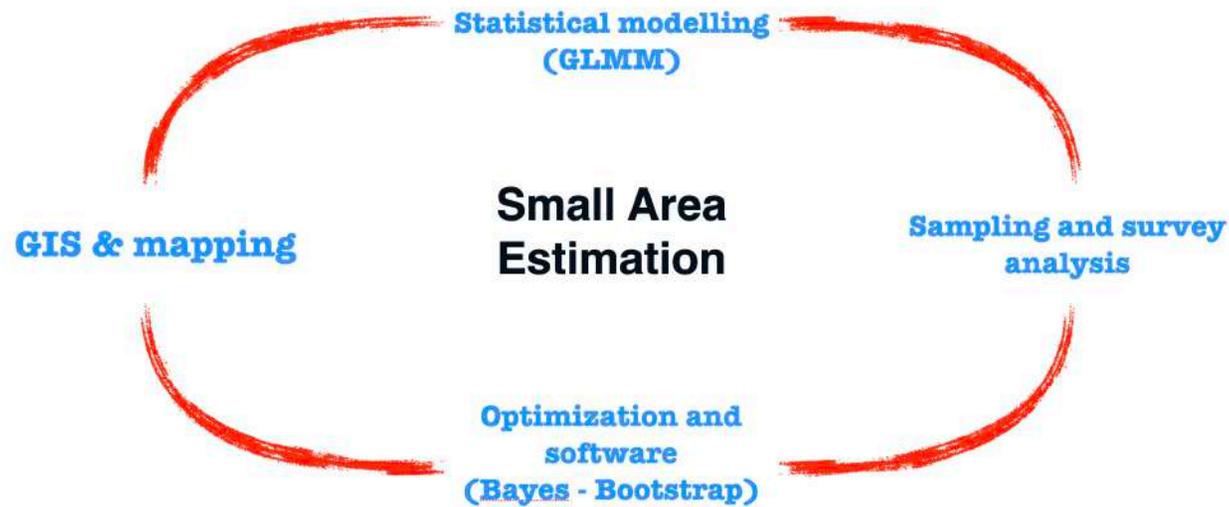
Benchmarking

- La CEPAL difunde cifras de pobreza nacional para los países de ALC con base en sus propias líneas de pobreza y datos de hogares de los países.
- Usamos un enfoque de *benchmarking* para crear ponderaciones poblacionales y obtener consistencia con las estimaciones agregadas provenientes de las encuestas de hogares.

Para ajustar las estimaciones de los modelos necesitamos encontrar un conjunto de ponderaciones g_k restringido a la siguiente condición:

$$\hat{R} = \frac{\sum_U g_k I_{y_k < z}}{\sum_U g_k} = R$$

Procesos en la producción de SAE



Fuente: adaptación de Kolenikov (2014).

Aplicaciones de CEPAL para la desagregación de indicadores de los ODS

Mapas de pobreza para países en Latinoamérica

Poner fin a la pobreza en todas sus formas en todo el mundo



- 1.1 Para 2030, erradicar la pobreza extrema para todas las personas en el mundo, actualmente medida por un ingreso por persona inferior a 1,25 dólares al día.
- 1.2 Para 2030, reducir al menos a la mitad la proporción de hombres, mujeres y niños y niñas de todas las edades que viven en la pobreza en todas sus dimensiones con arreglo a las definiciones nacionales.

Metodologías SAE para mapas de pobreza

Statistical Journal of the IAOS - Volume Pre-press, issue Pre-press



ISSN 1874-7655 (P)
ISSN 1875-9254 (E)

Poverty mapping in Latin America: ECLAC experiences on small area estimation

Cite

Authors: Gutiérrez, Andrés | Mancero, Xavier | Guerrero, Stalyn

Article Type: Research Article

Abstract: Poverty mapping is a valuable tool for governments and international organizations to identify the geographical areas and population groups that are most deprived or vulnerable. This approach can lead to designing and monitoring development policies more effectively. This article presents the recent experience of the Statistics Division of the UN Economic Commission for Latin America and the Caribbean (ECLAC) in using Small Area Estimation (SAE) methods to combine information from household surveys, censuses and satellite imagery to deliver poverty estimates at the provincial and municipal levels, which could not be attained using the household surveys alone.

Keywords: Poverty mapping, Pseudo-EBP, small area estimation, unit-level models

DOI: 10.3233/SJI-220037

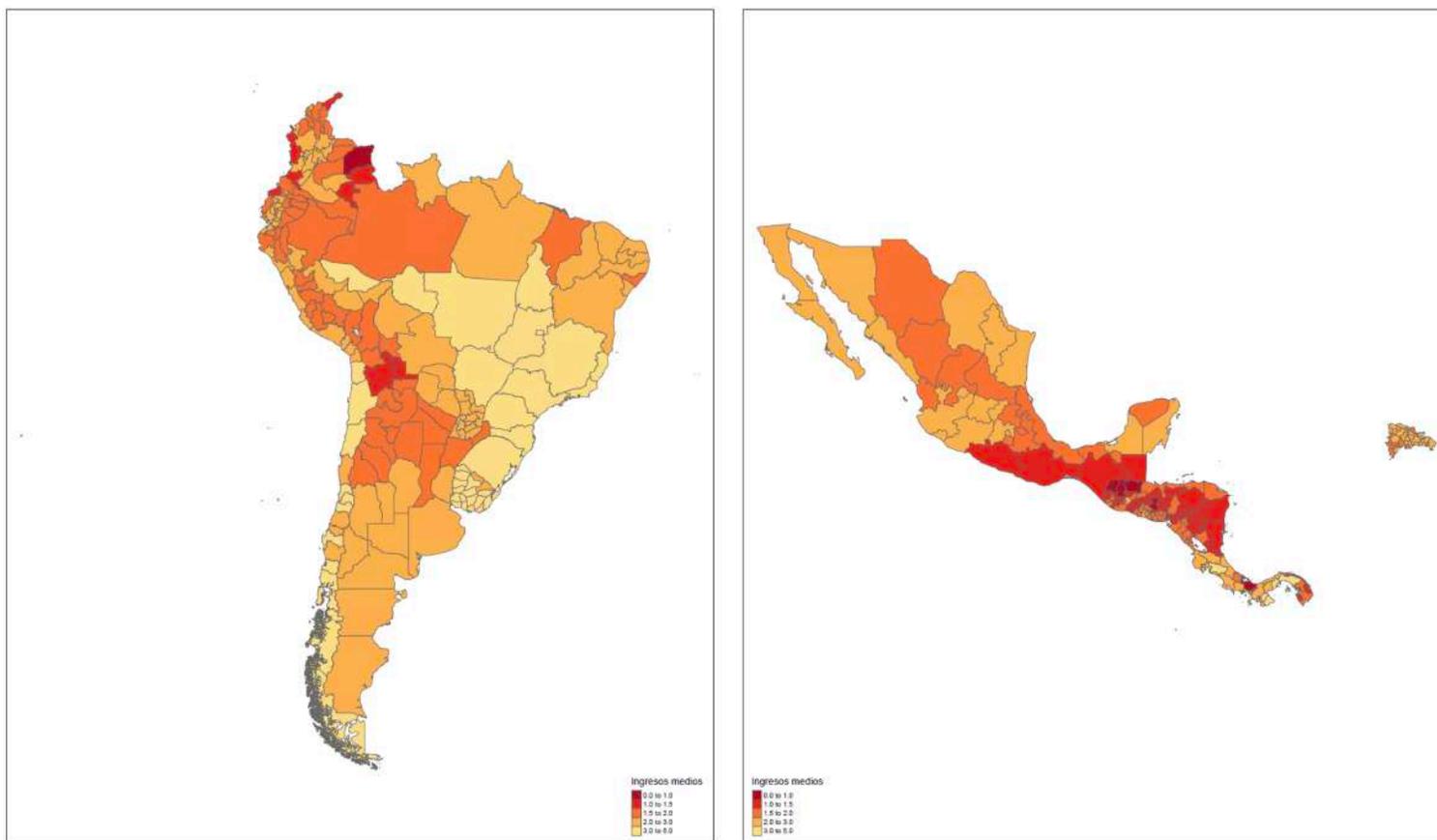
Citation: *Statistical Journal of the IAOS*, vol. Pre-press, no. Pre-press, pp. 1-13, 2022

Mapas de pobreza en la CEPAL

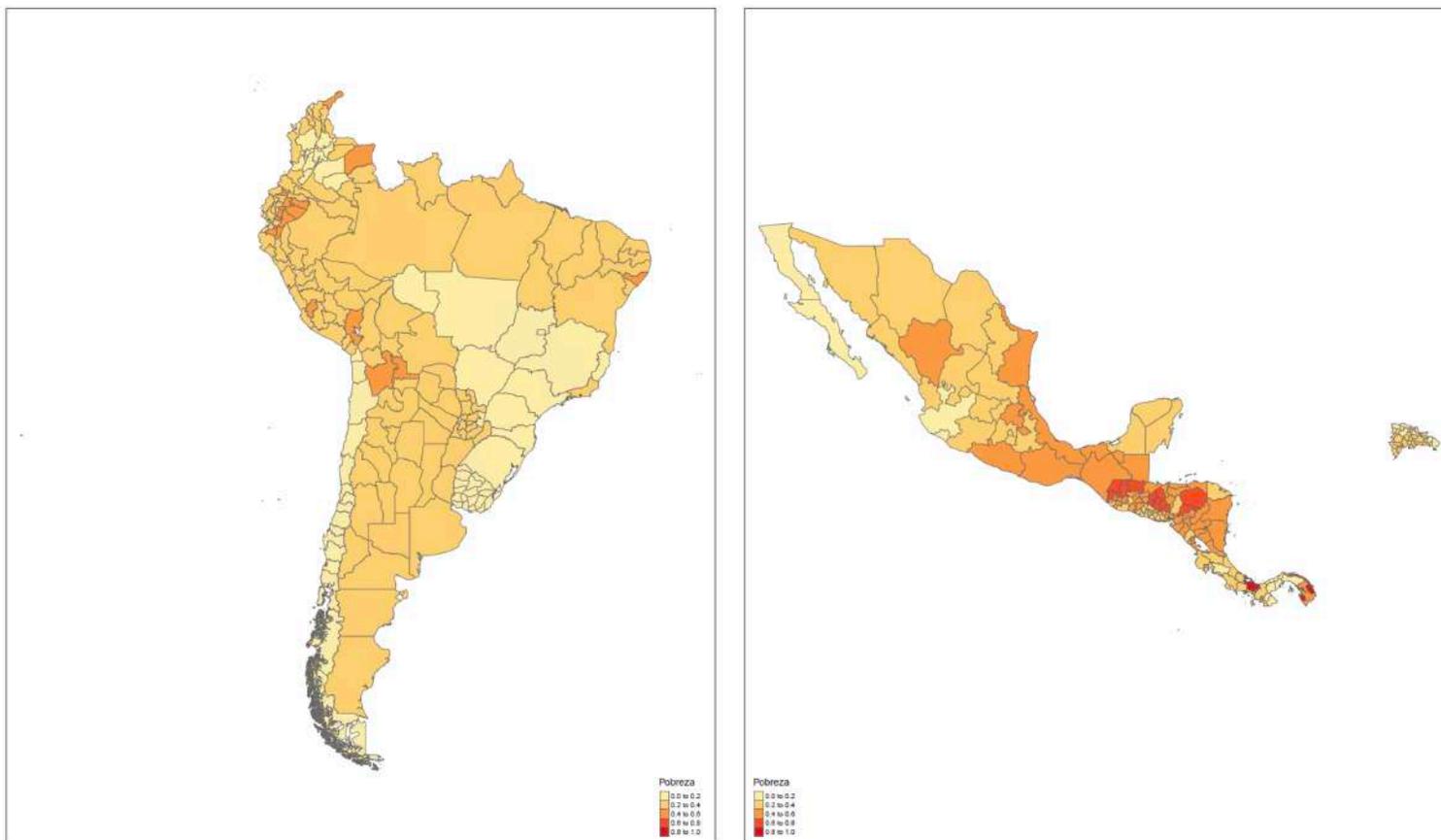
Hemos desarrollado un enfoque sistemático para producir estimaciones de SAE en los países de América Latina.

- Metodología actualizada.
- Cifras comparables en 18 países y cientos de dominios.
- Hasta donde sabemos, este sería el esfuerzo más extenso realizado actualmente para el mapeo de la pobreza en una gran región (América Latina).
- Actualizado anualmente, con encuestas continuas.
- Los conteos del censo se actualizan cada año para variables estructurales estables a lo largo de los años.

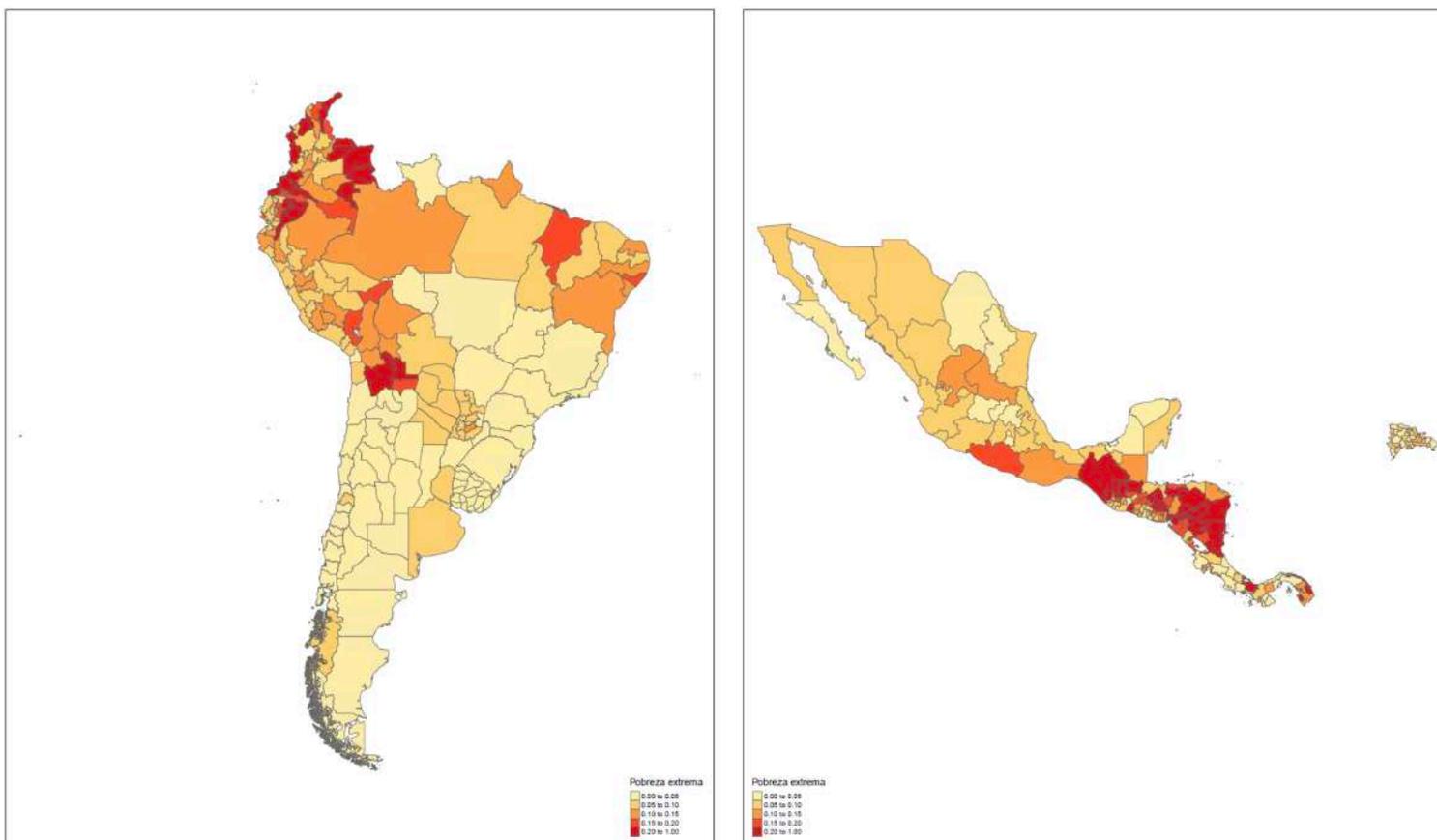
América Latina: estimaciones SAE a nivel estatal para el ingreso medio



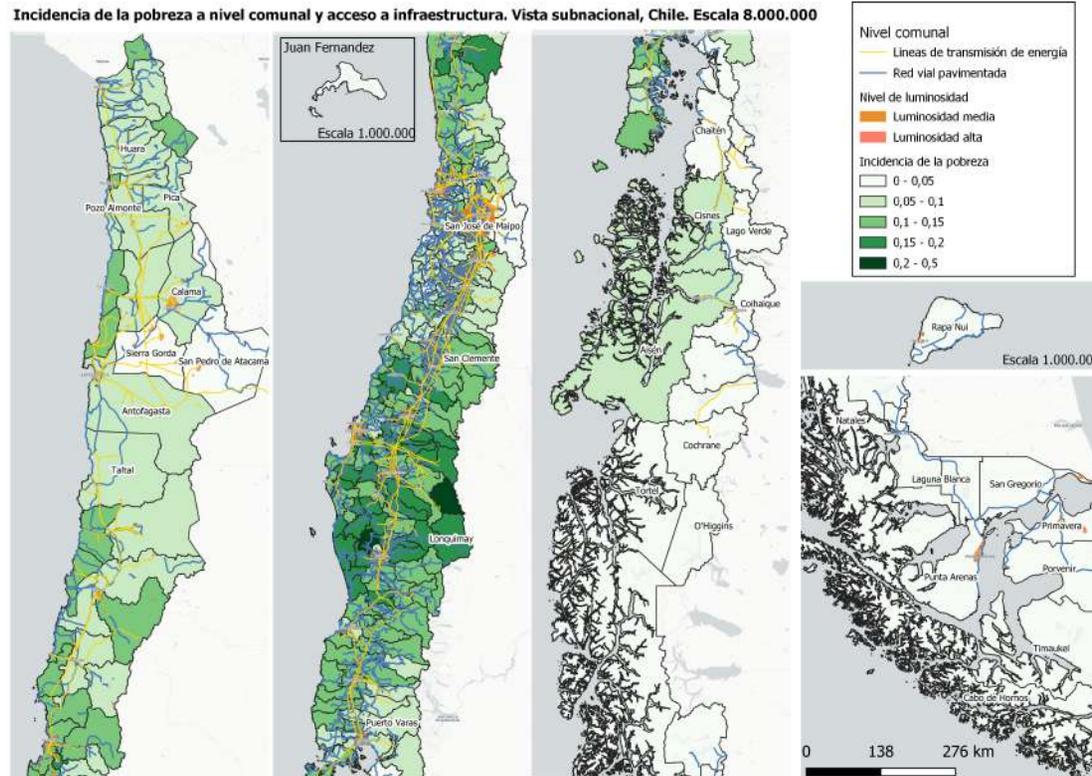
América Latina: estimaciones SAE a nivel estatal para la tasa de pobreza



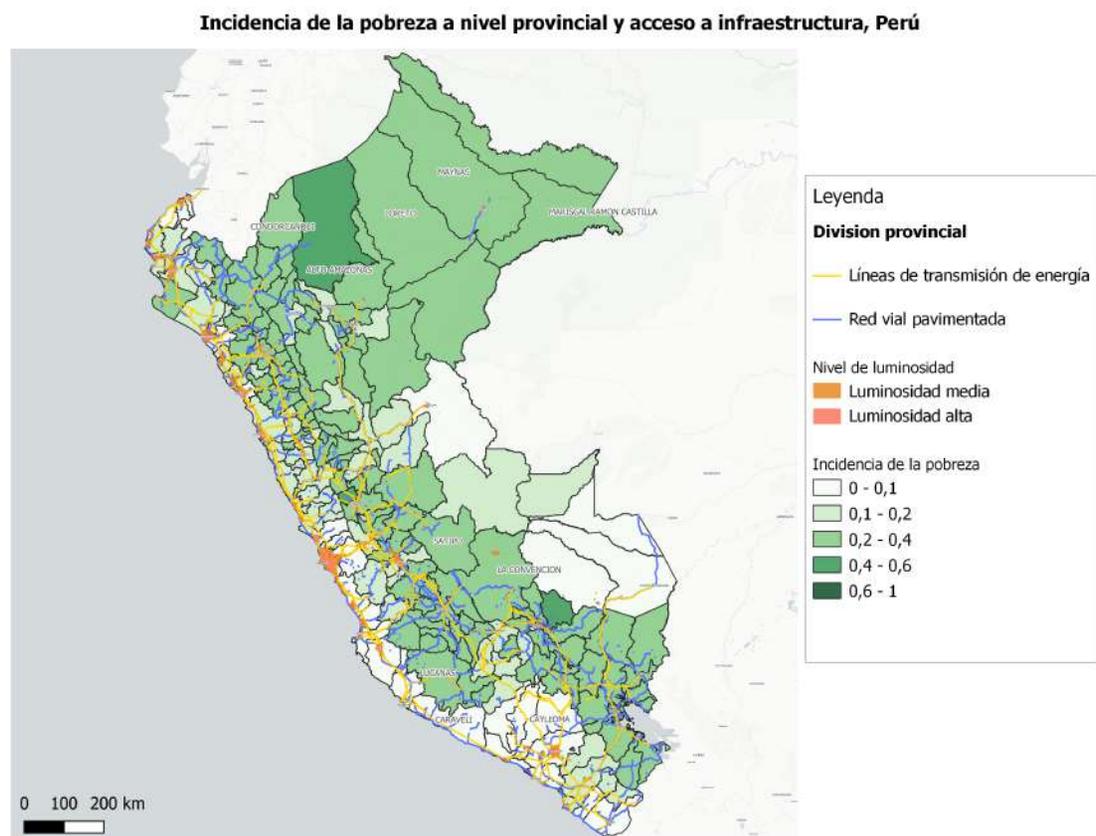
América Latina: estimaciones SAE a nivel estatal para la tasa de pobreza extrema



Chile: estimaciones SAE de pobreza monetaria a nivel de comuna



Perú: estimaciones SAE de pobreza monetaria a nivel provincial



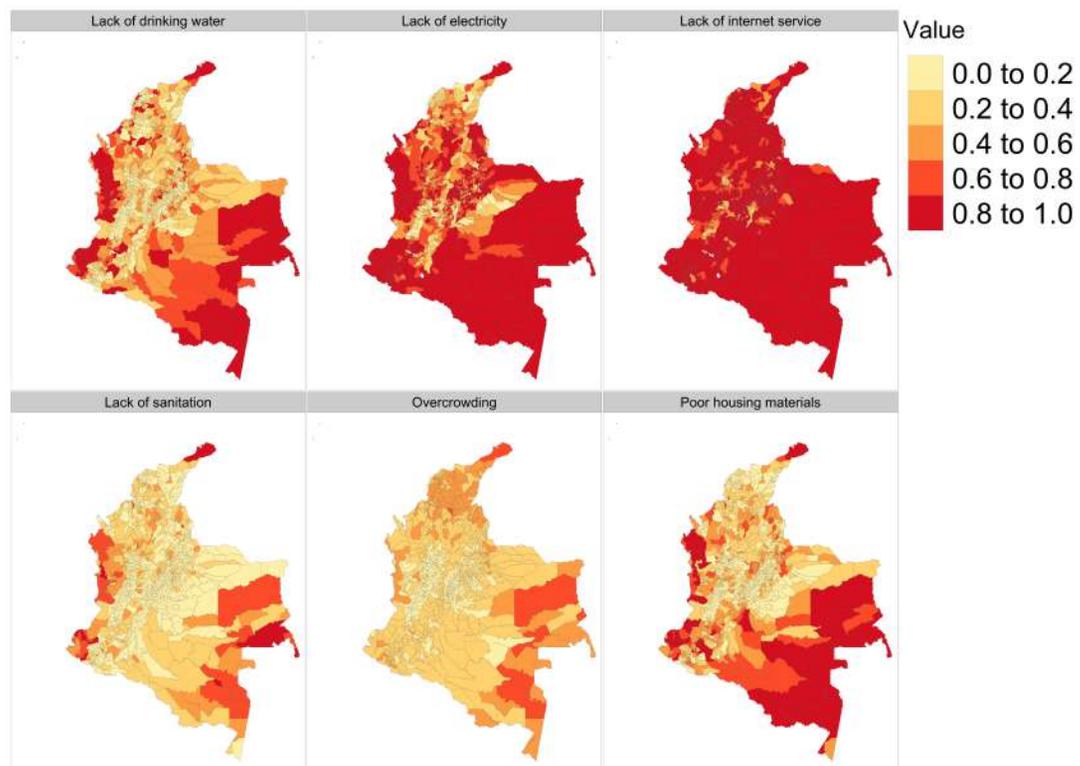
Privación multidimensional

Dimensiones del IPM de CEPAL (versión inicial)

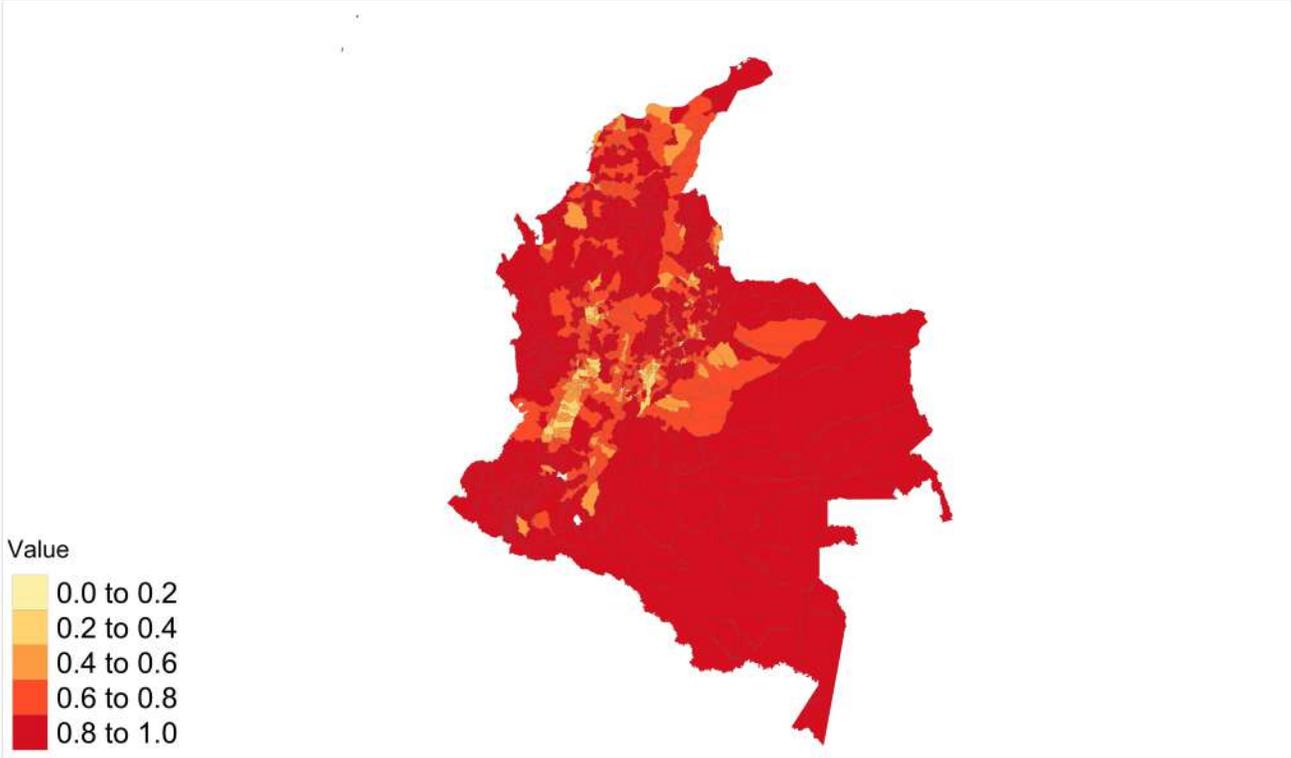
Table 1: Composition of the multidimensional deprivation index, availability of indicators in the Colombian census and target population.

Dimension	Indicator	Weight	Available in census	Target population
Housing	Poor housing materials	1/10	Yes	Adults and seniors
	Overcrowding	1/10	Yes	Adults and seniors
Water and sanitation	Lack of drinking water	1/10	Yes	Adults and seniors
	Lack of sanitation	1/10	Yes	Adults and seniors
Energy and connectivity	Lack of internet service	1/10	Yes	Adults and seniors
	Lack of electricity	1/10	Yes	Adults and seniors
Education	Unfinished education	2/10	No	Adults and Seniors
Employment and social protection	No or insufficient pension Unemployment or insufficient employment-related income	2/10	No	Seniors Adults

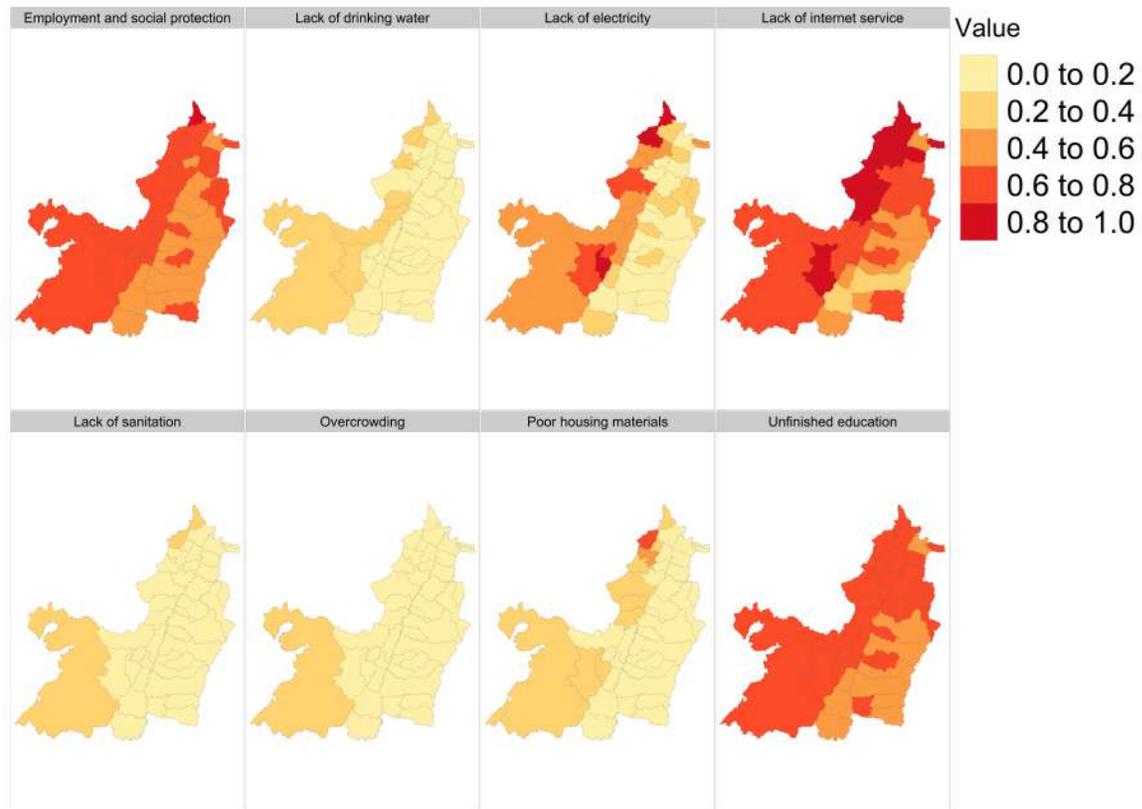
Colombia: dimensiones del IPM en el censo



Colombia: IPM final

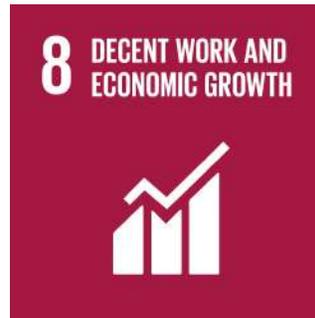


Valle del Cauca



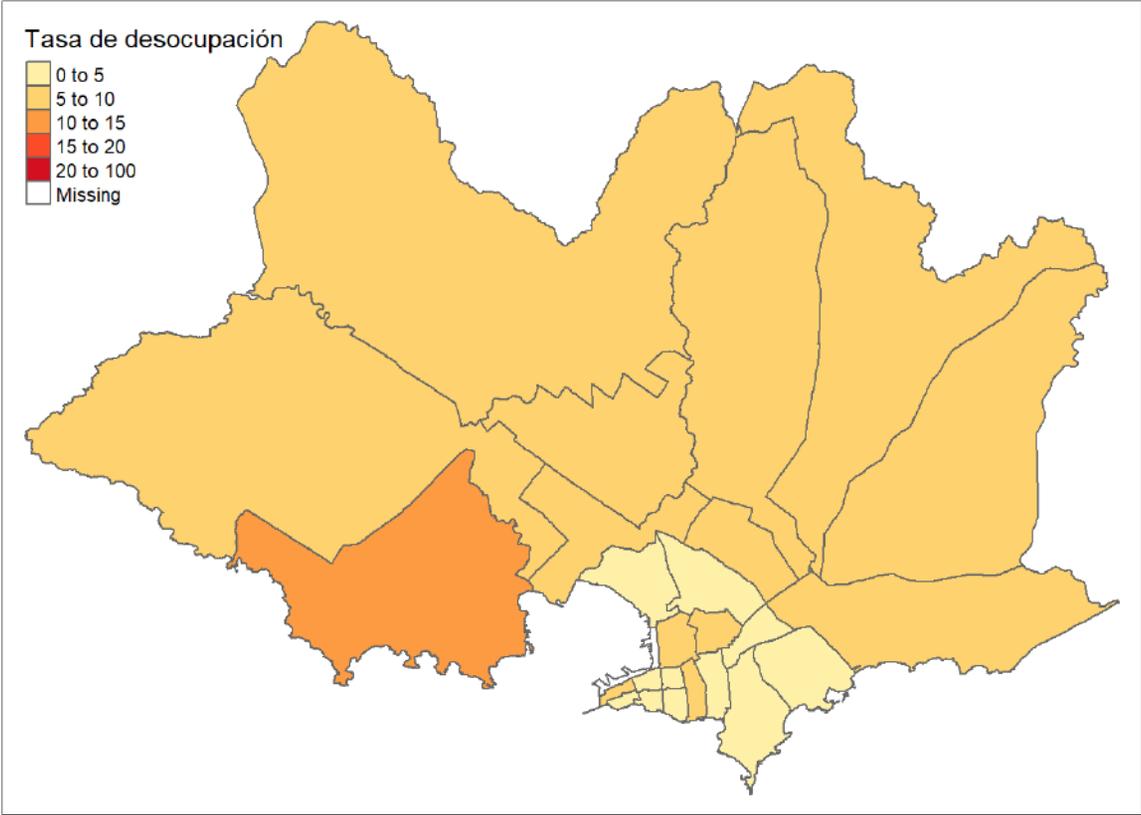
Estadísticas del mercado de trabajo

ODS 8: Trabajo decente y crecimiento económico

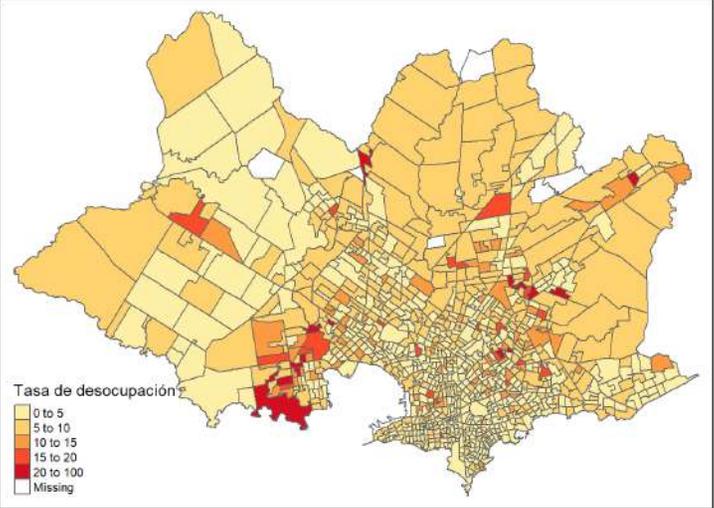
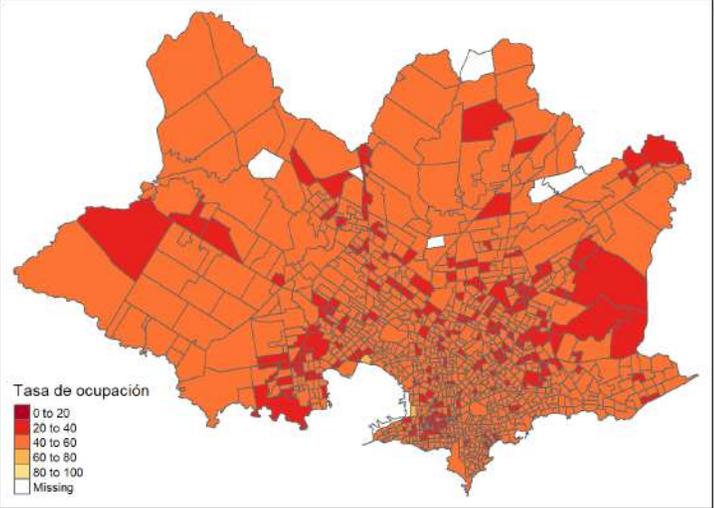


- 8.5 De aquí a 2030, lograr el empleo pleno y productivo y el trabajo decente para todas las mujeres y los hombres, incluidos los jóvenes y las personas con discapacidad, así como la igualdad de remuneración por trabajo de igual valor.
- 8.6 De aquí a 2020, reducir considerablemente la proporción de jóvenes que no están empleados y no cursan estudios ni reciben capacitación.

Montevideo: tasa de ocupación por segmento



Montevideo: tasa de ocupación y tasa de desocupación por sector



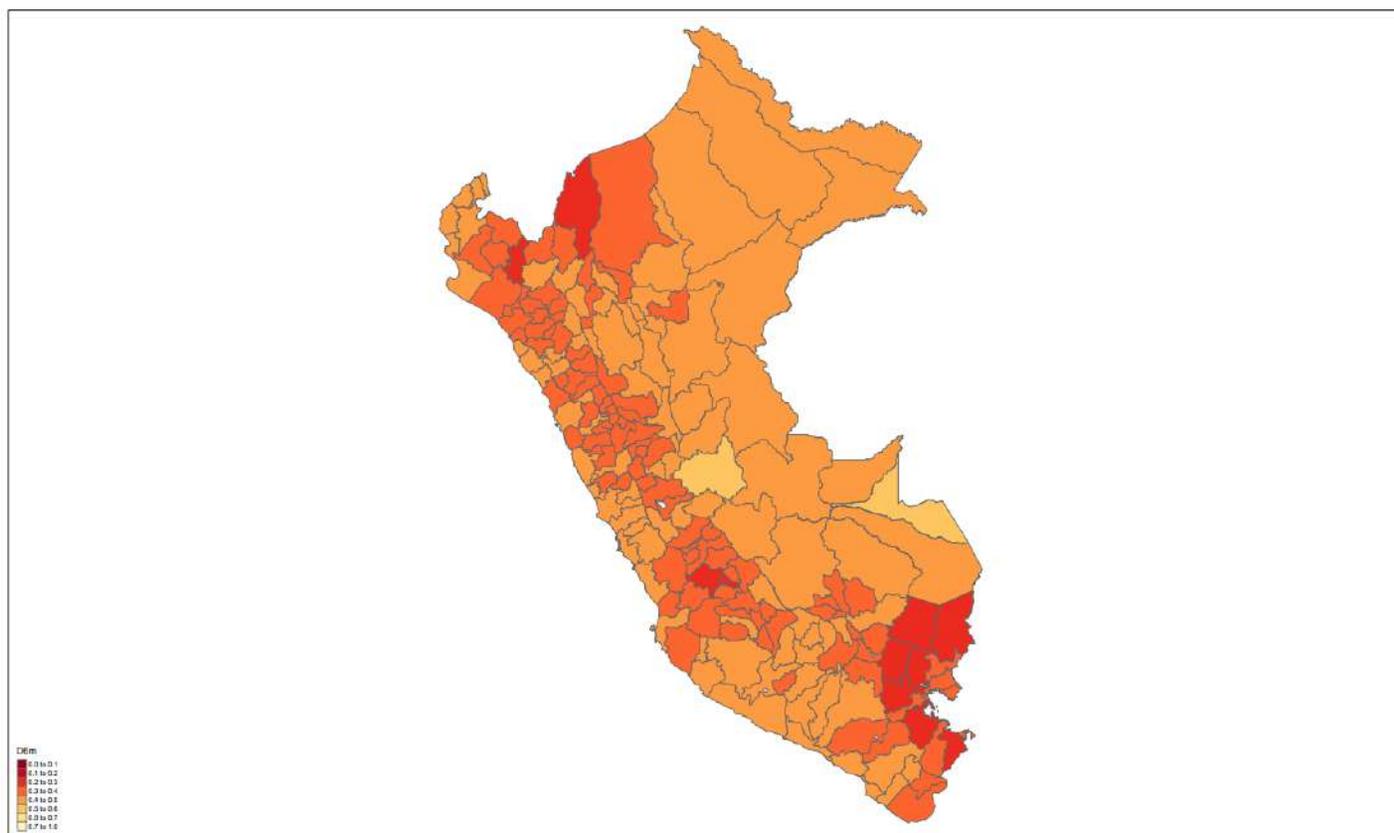
Métodos de planificación familiar

ODS 3: Salud y bienestar

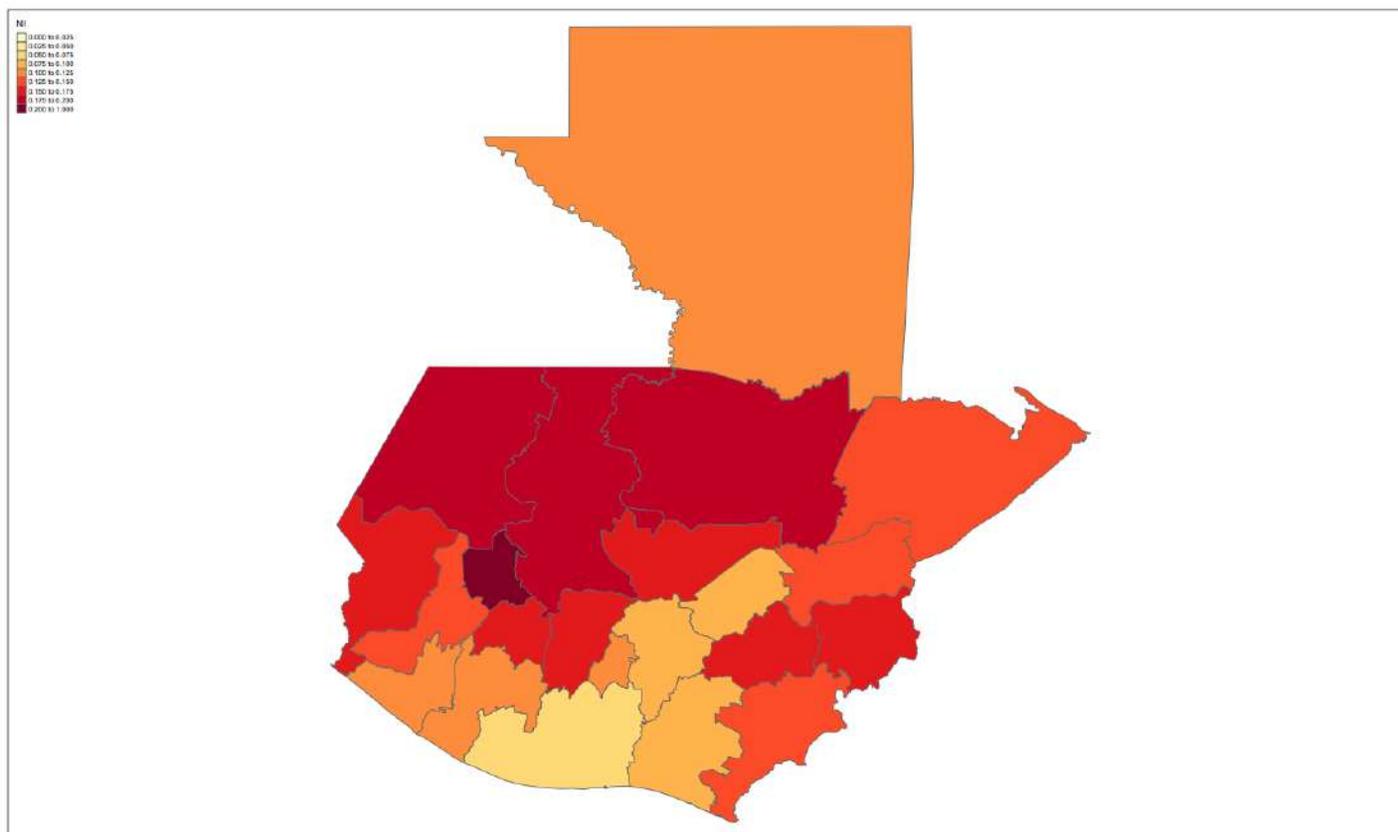


- 3.7 Para 2030, garantizar el acceso universal a los servicios de salud sexual y reproductiva, incluidos los de planificación de la familia, información y educación, y la integración de la salud reproductiva en las estrategias y los programas nacionales.
- 3.d Reforzar la capacidad de todos los países, en particular los países en desarrollo, en materia de alerta temprana, reducción de riesgos y gestión de los riesgos para la salud nacional y mundial.

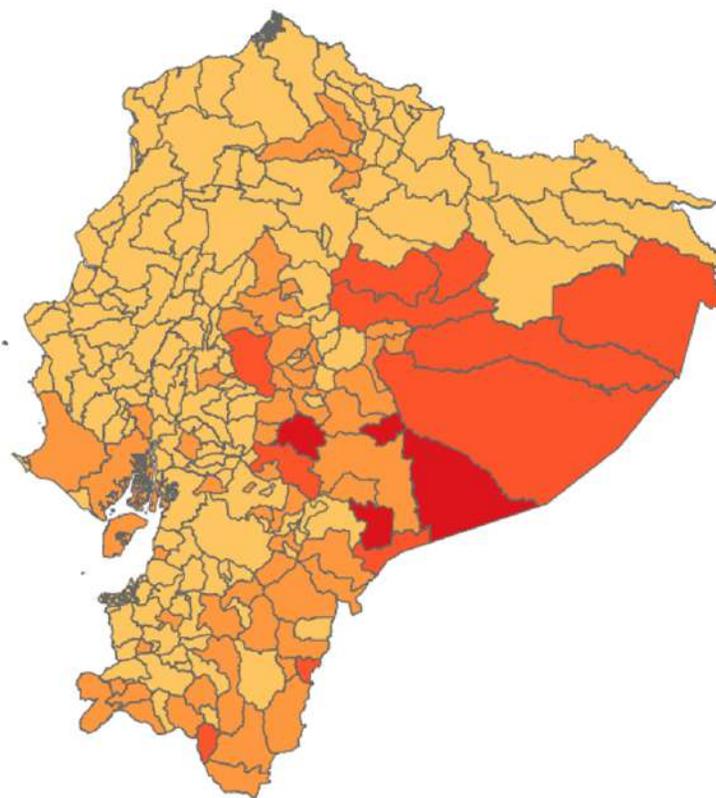
Perú: prevalencia del uso de métodos modernos de planificación familiar



Guatemala: porcentaje de mujeres con necesidades insatisfechas de planificación familiar



Ecuador: porcentaje de mujeres que cubren sus necesidades de planificación con métodos modernos



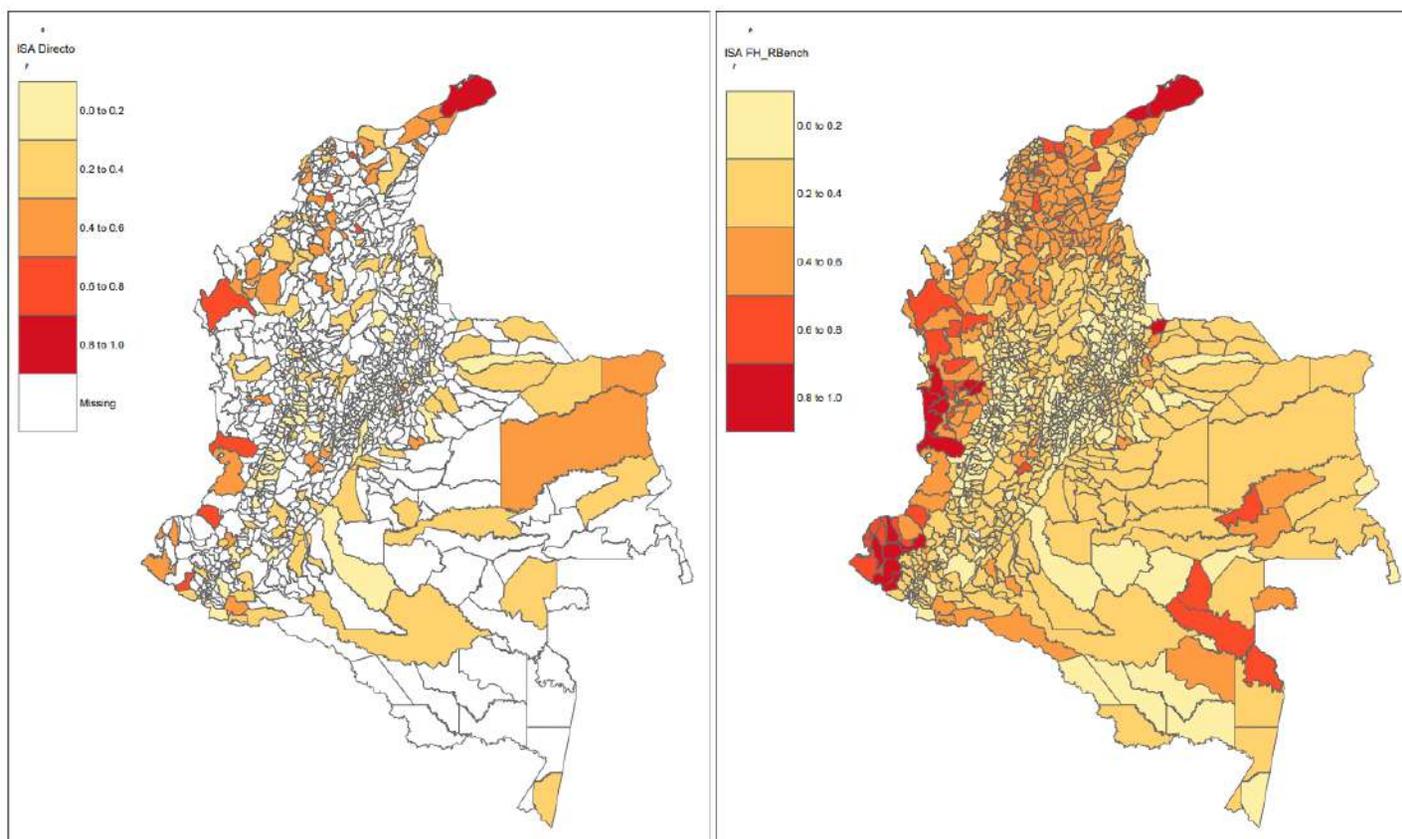
Inseguridad alimentaria

ODS 2: Hambre cero



- 2.1 Para 2030, poner fin al hambre y asegurar el acceso de todas las personas, en particular los pobres y las personas en situaciones vulnerables, incluidos los lactantes, a una alimentación sana, nutritiva y suficiente durante todo el año.
- 2.2 Para 2030, poner fin a todas las formas de malnutrición, incluso logrando, a más tardar en 2025, las metas convenidas internacionalmente sobre el retraso del crecimiento y la emaciación de los niños menores de 5 años, y abordar las necesidades de nutrición de las

Colombia: prevalencia de la inseguridad alimentaria a nivel municipal



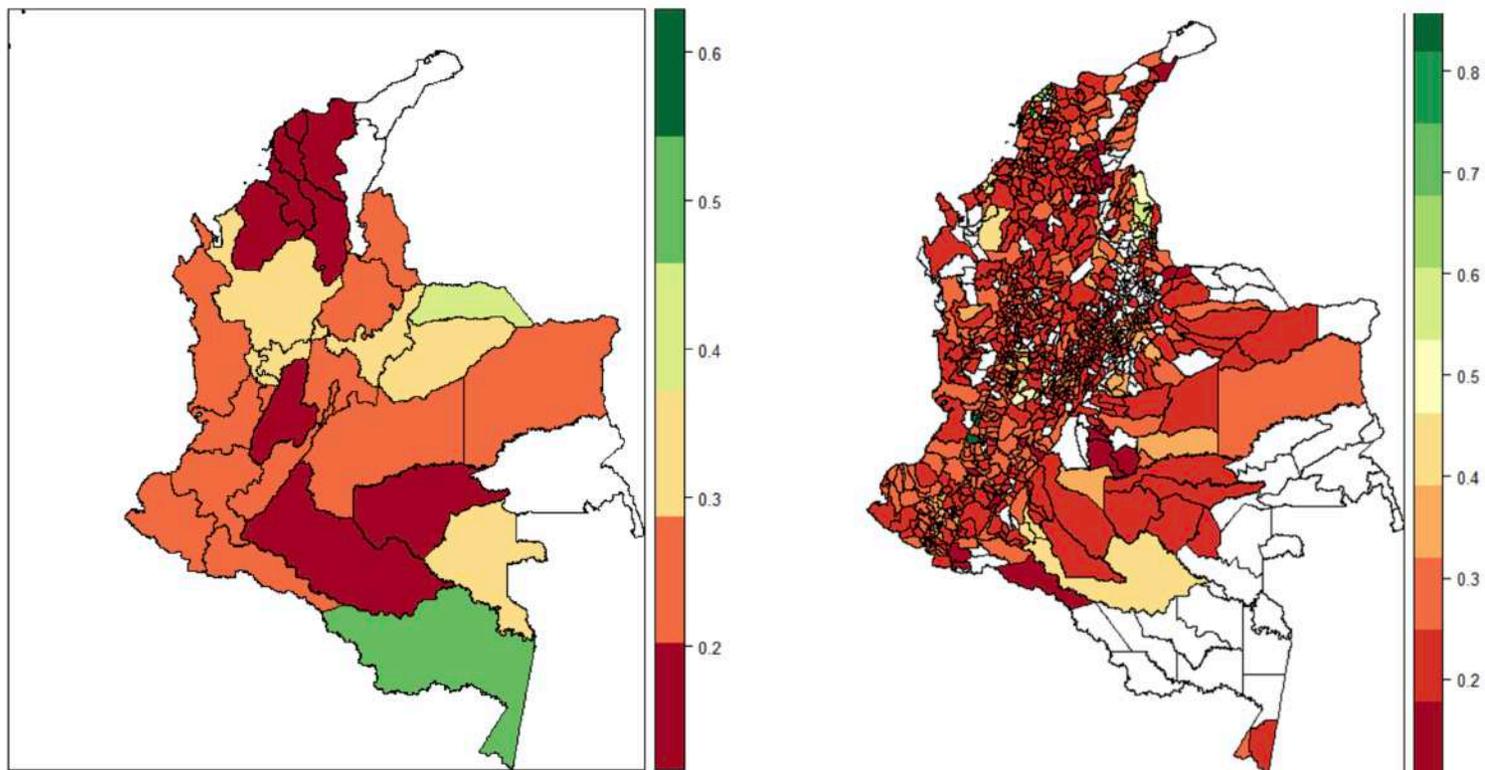
Acceso a la justicia

ODS 16: Paz, justicia e instituciones sólidas



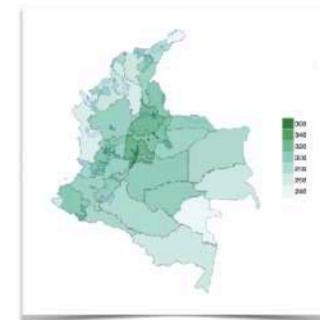
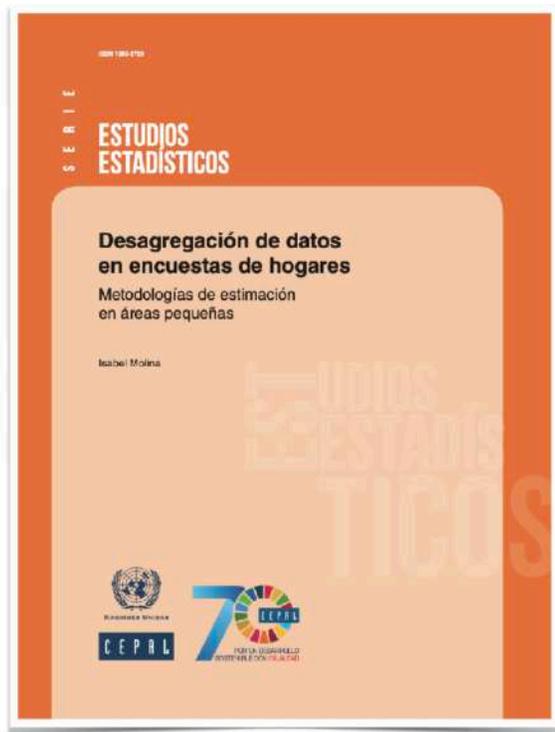
- 16.3 Promover el estado de derecho en los planos nacional e internacional y garantizar la igualdad de acceso a la justicia para todos.
- 16.10 Garantizar el acceso público a la información y proteger las libertades fundamentales, de conformidad con las leyes nacionales y los acuerdos internacionales.

Colombia: índice de acceso a la justicia a nivel departamental y municipal

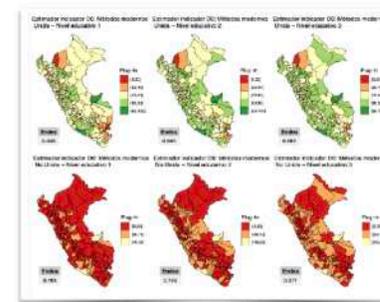


Entrenamiento en SAE por parte de CEPAL en la región

Entrenamiento



```
1> sigma <- 100
16 beta1 <- 0
17 beta0 <- 200
18 x1 <- runif(N, 0, 25)
19
20 y1 <- beta0 + 20 + beta1 * x1 + rnorm(N, 0, sigma)
21 y2 <- beta0 + 120 + beta1 * x1 + rnorm(N, 0, sigma)
22 y3 <- beta0 + 220 + beta1 * x1 + rnorm(N, 0, sigma)
23 y4 <- beta0 + 320 + beta1 * x1 + rnorm(N, 0, sigma)
24 y5 <- beta0 + 420 + beta1 * x1 + rnorm(N, 0, sigma)
25 ID <- rep(LETTERS[1:5], each = N)
26
27 test <- data.frame(Index = c(x1),
28                   Ingreso = c(y1, y2, y3, y4, y5), ID = ID)
29
30 p1 <- ggplot(data = test, aes(Index, Ingreso)) +
31   geom_point() + geom_smooth(method='lm') + theme_bw()
32
```



Desarrollo de un sistema de estimación SAE en ALC

Gracias!

Rolando Ocampo

Xavier Mancero

Andrés Gutiérrez