

Desagregación de indicadores ODS: el enfoque cepalino

**Andrés Gutiérrez
División de Estadísticas
CEPAL**

2024



SUSTAINABLE DEVELOPMENT GOALS

1 NO POVERTY

2 ZERO HUNGER

3 GOOD HEALTH AND WELL-BEING

4 QUALITY EDUCATION

5 GENDER EQUALITY

6 CLEAN WATER AND SANITATION

7 AFFORDABLE AND CLEAN ENERGY

8 DECENT WORK AND ECONOMIC GROWTH

9 INDUSTRY, INNOVATION AND INFRASTRUCTURE

10 REDUCED INEQUALITIES

11 SUSTAINABLE CITIES AND COMMUNITIES

12 RESPONSIBLE CONSUMPTION AND PRODUCTION

13 CLIMATE ACTION

14 LIFE BELOW WATER

15 LIFE ON LAND

16 PEACE, JUSTICE AND STRONG INSTITUTIONS

17 PARTNERSHIPS FOR THE GOALS

SUSTAINABLE DEVELOPMENT GOALS

Los ODS se centran en la desagregación

Objetivo 1.2. Para 2030, reducir al menos a la mitad la proporción de hombres, mujeres y niños de todas las edades que viven en la pobreza en todas sus dimensiones, de acuerdo con las definiciones nacionales.

- Indicador 1.2.2: Proporción de hombres, mujeres y niños de todas las edades que viven en la pobreza en todas sus dimensiones, según las definiciones nacionales.

Meta 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, y la igualdad de remuneración por trabajo de igual valor.

- Indicador 8.5.1: Ingresos medios por hora de los asalariados, por sexo, edad, ocupación y personas con discapacidad.



**LEAVE NO
ONE BEHIND**



Principio fundamental de la desagregación de datos

Los indicadores de los Objetivos de Desarrollo Sostenible deben desglosarse, cuando proceda, por ingresos, sexo, edad, raza, etnia, situación migratoria, discapacidad y ubicación geográfica, u otras características, de conformidad con los Principios Fundamentales de las Estadísticas Oficiales.

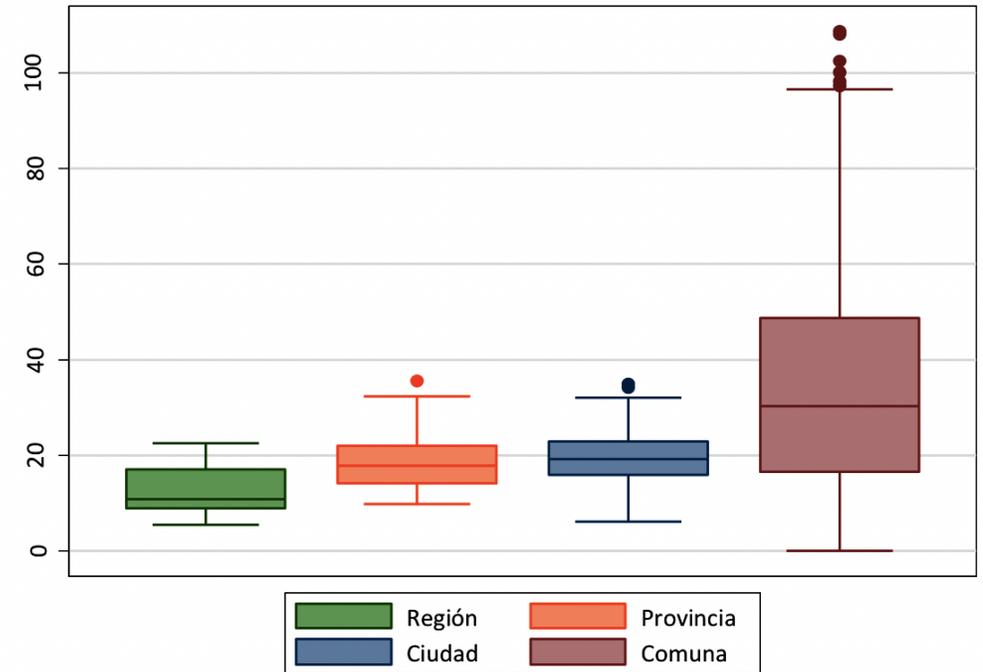
**Resolución de la Asamblea General -
68/261**

Los modelos SAE en la CEPAL

Limitaciones de las encuestas

Las bondades de las encuestas dependen de un tamaño de muestra grande. Cuando las condiciones son apropiadas, proporcionan un sólido sistema inferencial que proporciona una estimación precisa y exacta en dominios planificados.

Pero, cuando no...



El problema

Cuando el tamaño de la muestra no permite obtener estimaciones directas confiables para algunos dominios de interés, se pueden abordar las siguientes opciones:

1. Aumentar el tamaño de la muestra: esta opción eleva los costos y es inviable.
2. Utilizar metodologías estadísticas que involucren información auxiliar externa para obtener estimaciones confiables en los subgrupos de interés, manteniendo el tamaño de la muestra de la encuesta..

Tipos de áreas pequeñas

Es necesario recurrir a información auxiliar externa (censos, registros administrativos, imágenes satelitales) para que en conjunto (encuestas y datos externos) se pueda construir un sistema inferencial preciso.

- Las unidades geográficas pueden incluir estados, provincias, departamentos o municipios.
- Los subgrupos específicos pueden incluir combinaciones como la edad × el sexo × la etnia × el estado migratorio.

La CEPAL utiliza modelos de SAE para integrar datos de diferentes fuentes.

Enfoque de la CEPAL

La CEPAL usa una metodología estandarizada que sigue las siguientes etapas:

- Etapa 1. Estandarización y homologación de covariables en las bases de datos (censos y encuestas de hogares). Adaptación de imágenes satelitales como covariables a nivel estatal o municipal.
- Etapa 2. Actualización de los conteos intercensales relacionados con las covariables, preservando las estructuras censales y actualizando los márgenes de la encuesta de hogares.
- Etapa 3. Definición de modelos bayesianos considerando posibles interacciones, selección de variables auxiliares y estimación de los coeficientes del modelo.

Enfoque de la CEPAL

- Etapa 4. Predicción de indicadores sobre postestratos censales y áreas pequeñas, estimación del MSE a partir de réplicas MCMC.
- Etapa 5: Validación de los supuestos del modelo y evaluación comparativa utilizando estimaciones directas a nivel de los dominios de representatividad de las encuestas.
- Etapa 6: Generación de mapas para los países de América Latina.

Modelo SAE para pobreza

Se sigue el enfoque de Molina y Rao (2010), quienes estimaron indicadores de pobreza utilizando el siguiente modelo (EBLUP).

Modelo de vínculo:

$$\tilde{\delta}_d^B(\boldsymbol{\theta}) = E_{\mathbf{y}_{dr}}[\delta_d(\mathbf{y}_d) | \mathbf{y}_{ds}; \boldsymbol{\theta}]$$
$$\mathbf{y}_{dr} | \mathbf{y}_{ds} \sim_{ind} N(\boldsymbol{\mu}_{dr|s}, \mathbf{V}_{dr|s})$$

Estimación condicional:

$$\boldsymbol{\mu}_{dr|s} = \mathbf{X}_{dr}\boldsymbol{\beta} + \gamma_d(\bar{y}_{da} - \bar{\mathbf{x}}_{da}^T\boldsymbol{\beta})\mathbf{1}_{N_d-n_d}$$
$$\mathbf{V}_{dr|s} = \sigma_u^2(1 - \gamma_d)\mathbf{1}_{N_d-n_d}\mathbf{1}_{N_d-n_d}^T + \sigma_e^2 \text{diag}_{i \in r_d}(k_{di}^2)$$

Modelo SAE para indicadores de fuerza laboral

Seguimos el enfoque de Molina, Saei y Lombardía (2007), quienes propusieron una metodología basada en el supuesto de que los totales de la muestra de individuos siguen un modelo logit multinomial con efectos aleatorios de área.

Se considera un vector multinomial que cuenta el número de individuos desempleados (y_1), empleados (y_2) e inactivos (y_3) muestreados dentro de cada área pequeña.

$$\log \left(\frac{p_{dij}}{p_{di3}} \right) = \mathbf{x}_{di} \boldsymbol{\beta}_j + u_d$$

Integración de datos para SAE

Banco de Datos de Encuestas de Hogares de la CEPAL (BADEHOG)

- Se trata de un repositorio de encuestas de hogares de 18 países de América Latina mantenido por la División de Estadística.
- Es un insumo esencial para los indicadores sociales armonizados elaborados por la CEPAL y para apoyar las investigaciones y los estudios de la CEPAL y algunos organismos de las Naciones Unidas.
- Anualmente recopila datos de encuestas de hogares de 18 países de América Latina.
- Se tienen encuestas a partir de 1980. En general, se tiene una encuesta por año por país.

País	Encuesta	
ARG	Encuesta Permanente de Hogares (EPH)	
BOL	Encuesta Nacional de Hogares	
BRA	Encuesta Nacional por Muestra Continua de Hogares	
CHL	Encuesta Nacional de Caracterización Socioeconómica (CASEN)	
COL	Gran Encuesta Integrada de Hogares	
CRI	Encuesta Nacional de Hogares (ENAHO)	
DOM	Encuesta Nacional Continua de Fuerza de Trabajo (ENCFT)	
ECU	Encuesta Nacional de Empleo, Desocupación y Subempleo (ENEMDU)	
GTM	Encuesta Nacional de Condiciones de Vida	
HND	Encuesta Multipropósito de Hogares	
MEX	Encuesta Nacional de Ingresos y Gastos de los Hogares (ENIGH)	
NIC	Encuesta Nacional de Hogares para la Medición del Nivel de Vida	
PAN	Encuesta multipropósito	
PER	Encuesta Nacional de Hogares - Condiciones de Vida y Pobreza	
PRY	Encuesta Permanente Continua de Hogares (EPHC)	
SLV	Encuesta de Hogares de Propósitos Múltiples	
URY	Encuesta Continua de Hogares	

Banco de datos censales de CELADE

- Se trata de un repositorio de censos de América Latina y el Caribe mantenido por la División de Población de la CEPAL (CELADE), que ha continuado con su actividad permanente de difusión de la información censal derivada de su banco de datos.
- La mayoría de los países latinoamericanos han compartido sus bases de datos con CELADE, lo que permite la disponibilidad de microdatos de las rondas de censos anteriores y actuales.
- Para nuestras necesidades de procesamiento estadístico empleamos el software Redatam, desarrollado por CELADE. Este software está diseñado específicamente para el procesamiento de microdatos relacionados con los censos de población y vivienda.

País	Censo	
ARG	Censo Nacional de Población, Hogares y Vivienda	
BOL	Censo de Población y Vivienda	
BRA	Censo Demográfico	
CHL	Censo de Población y Vivienda	
COL	Censo Nacional de Población y Vivienda (CNPV)	
CRI	X Censo Nacional de Población y Vivienda	
DOM	IX Censo Nacional de Población y Vivienda	
ECU	VII Censo de Población y Vivienda	
GTM	XII Censo Nacional de Población y VII Censo de Vivienda	
HND	XVII Censo de Población y Vivienda	
MEX	Censo de Población y Vivienda	
NIC	Censo de Población y Vivienda de Nicaragua	
PAN	Censo 2010	
PER	XII Censo de Población, VII Censo de Vivienda y III Censo de Comunidades Indígenas	
PRY	II Censo Nacional Indígena de Población y Vivienda	
SLV	VI Censo de Población y V Censo de Vivienda	
URY	Censo de población	

Imágenes satelitales

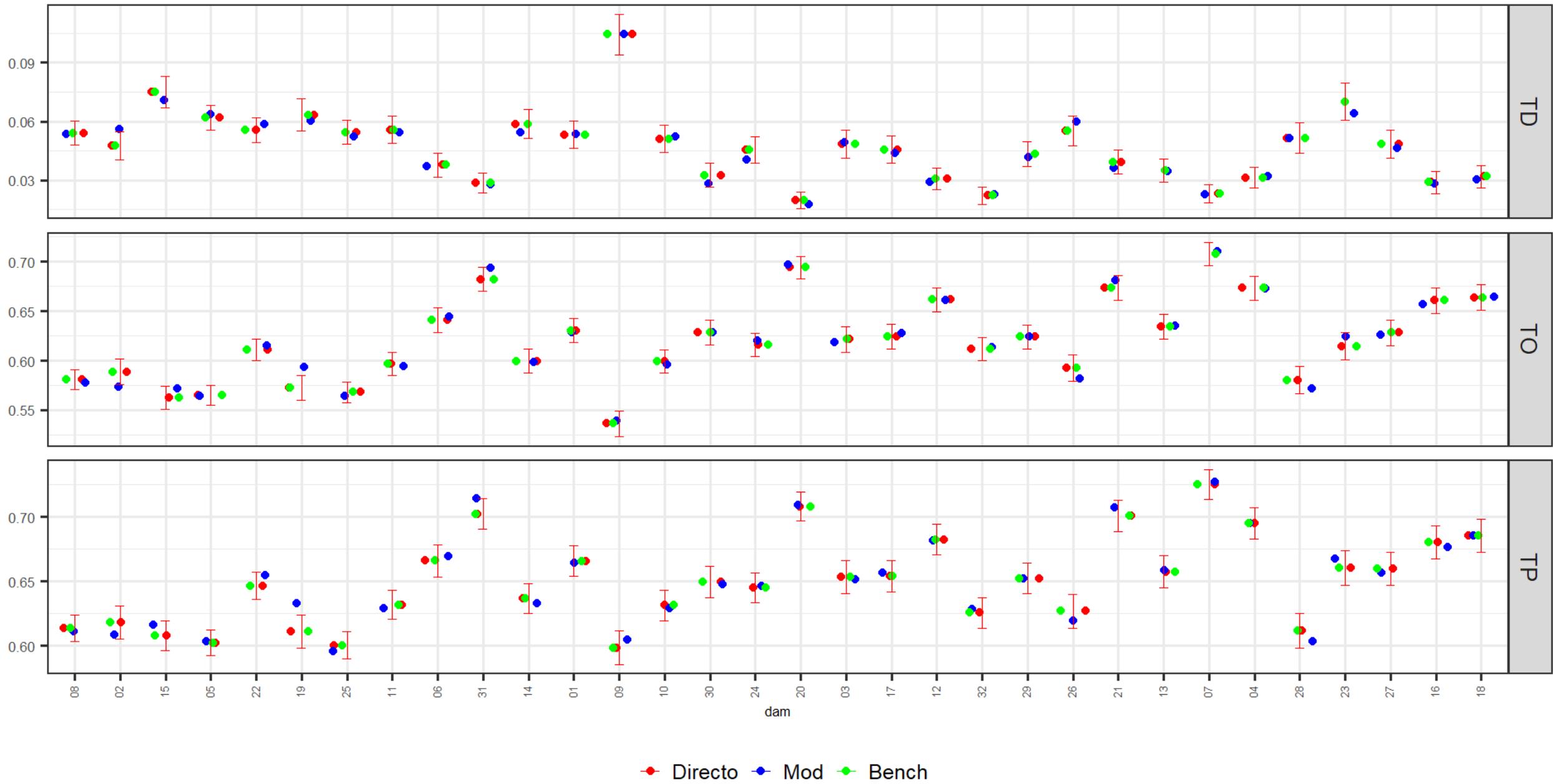
Entre las principales ventajas de la información basada en teledetección se encuentra la facilidad de acceso a datos con una cobertura geográfica profunda que es imposible obtener por medios tradicionales como encuestas o registros administrativos.

- Accedemos a esta información a través de Google Earth Engine, que proporciona facilidades para analizar y obtener estos datos a través de los lenguajes de programación Javascript y Python, y recientemente en R con el paquete rgee.
- La División de Estadística de la CEPAL está elaborando el Proyecto de Banco de Imágenes de Productos del Procesamiento de Imágenes por Satélite, destinado a procesar datos de teleobservación para utilizarlos en diversos proyectos, incluidas las estimaciones de la pobreza y el desempleo.

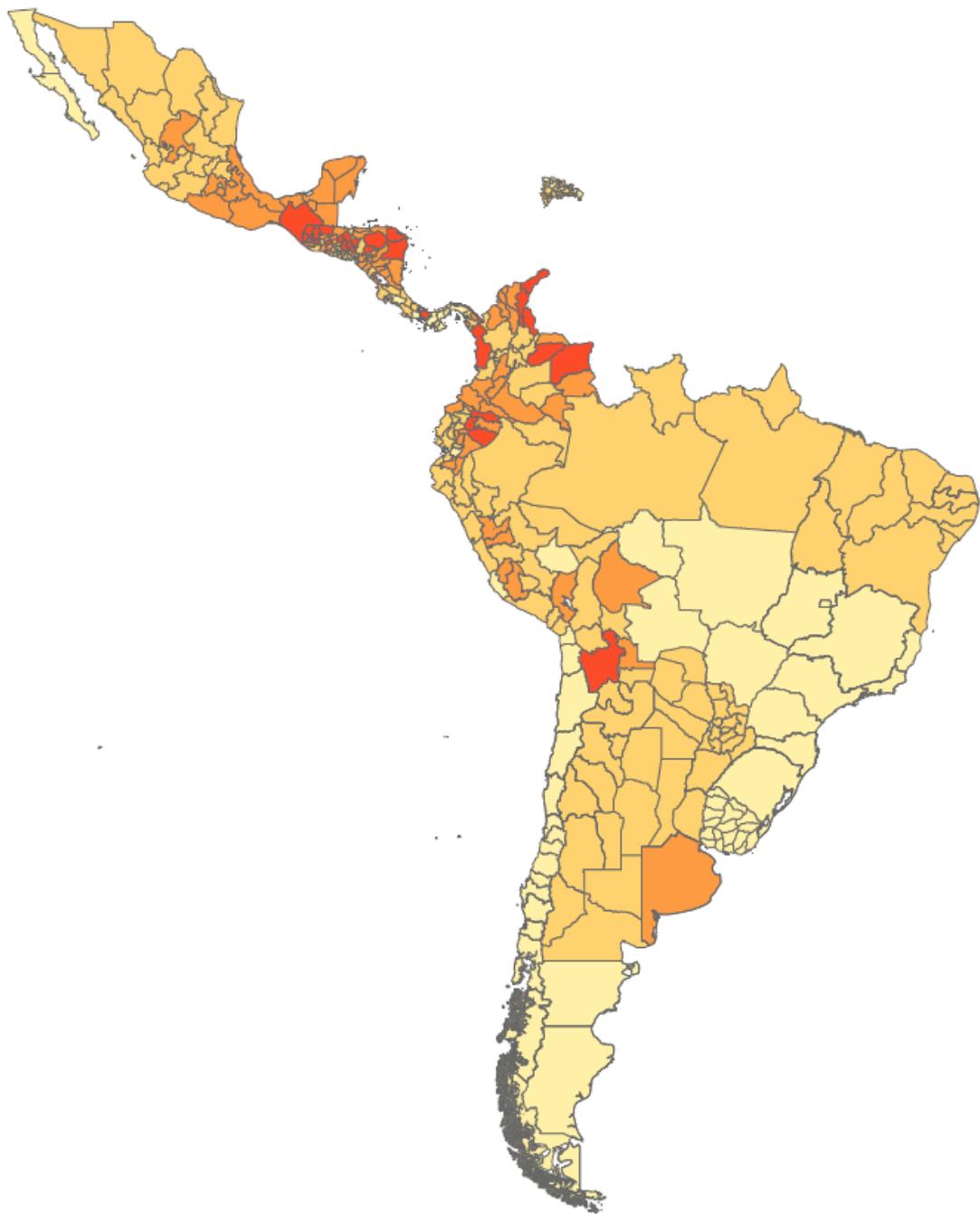
Benchmarking y evaluación de modelos

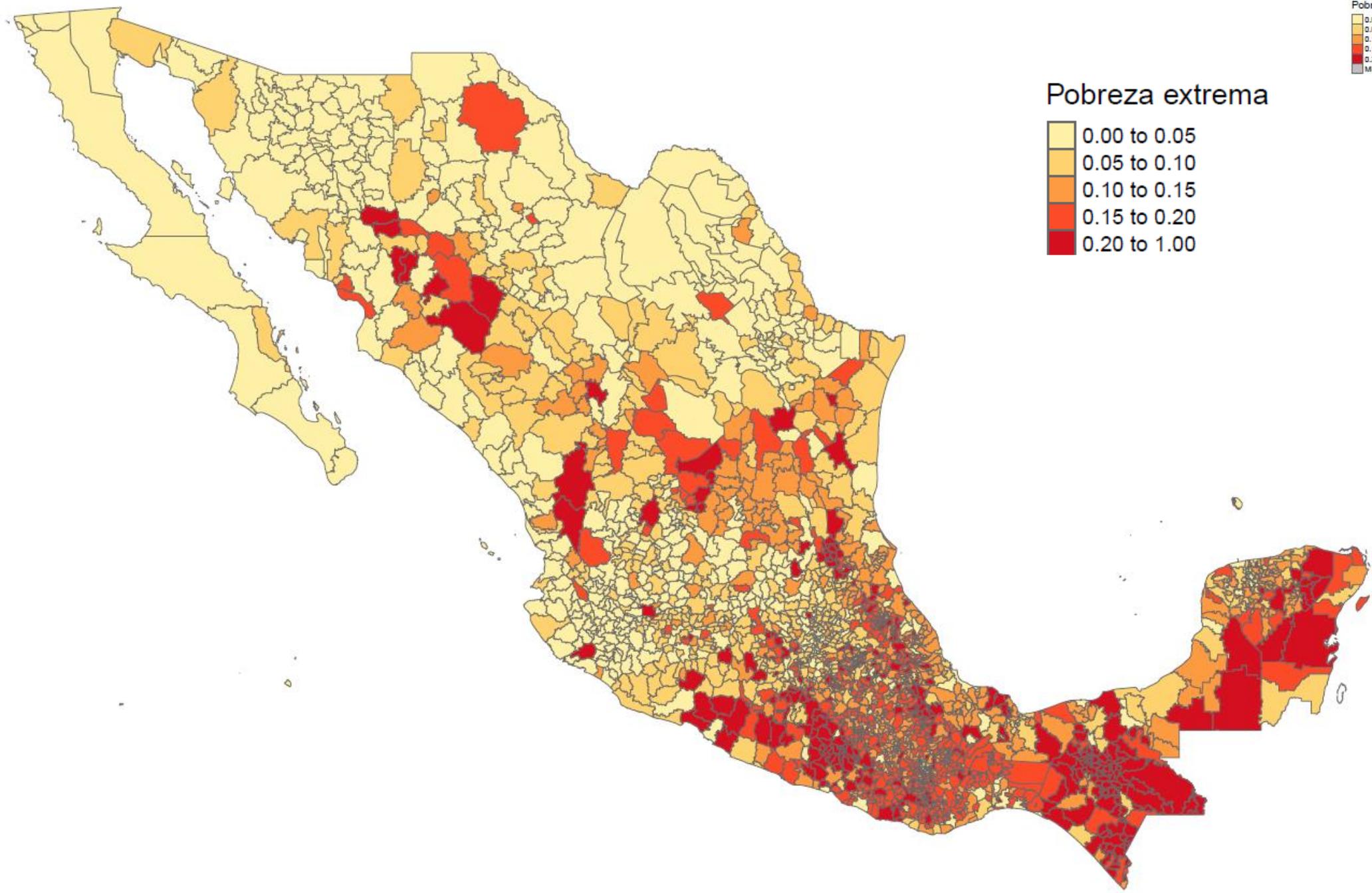
Una vez que las cadenas han alcanzado la convergencia, utilizamos las estimaciones directas a nivel nacional y regional para comparar las estimaciones de áreas pequeñas. Este proceso se lleva a cabo para cada iteración de un proceso de monte Carlo de Markov Chain.

Llevamos a cabo un estricto proceso de validación y evaluación del modelo para detectar posibles violaciones de los supuestos del modelo. Las pruebas utilizadas evalúan la normalidad y la heterocedasticidad. Además, utilizamos las distancias de Cook como medidas de la influencia de las observaciones en la muestra.



Maps for poverty indicators





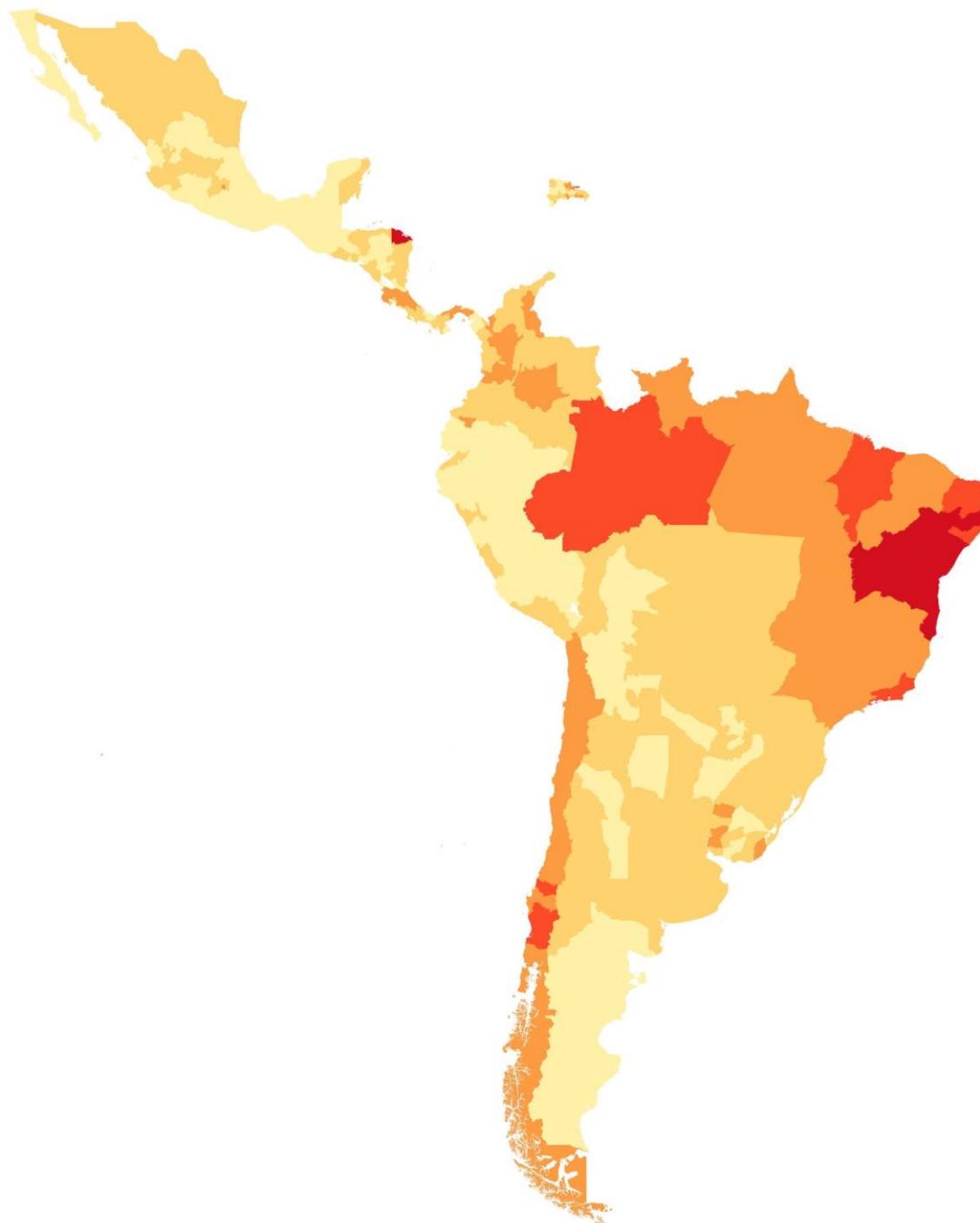
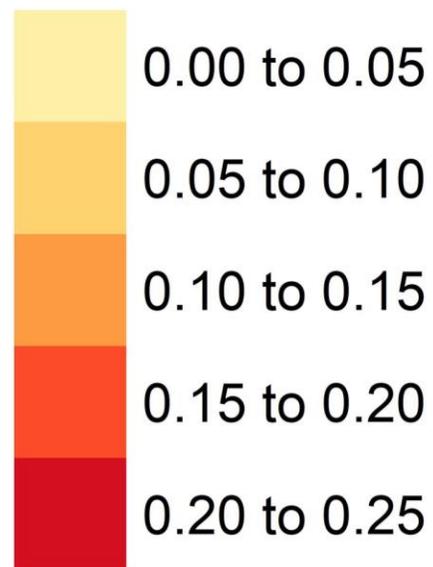
Pobreza extrema
0.00 to 0.05
0.05 to 0.10
0.10 to 0.15
0.15 to 0.20
0.20 to 1.00
Missing

Pobreza extrema

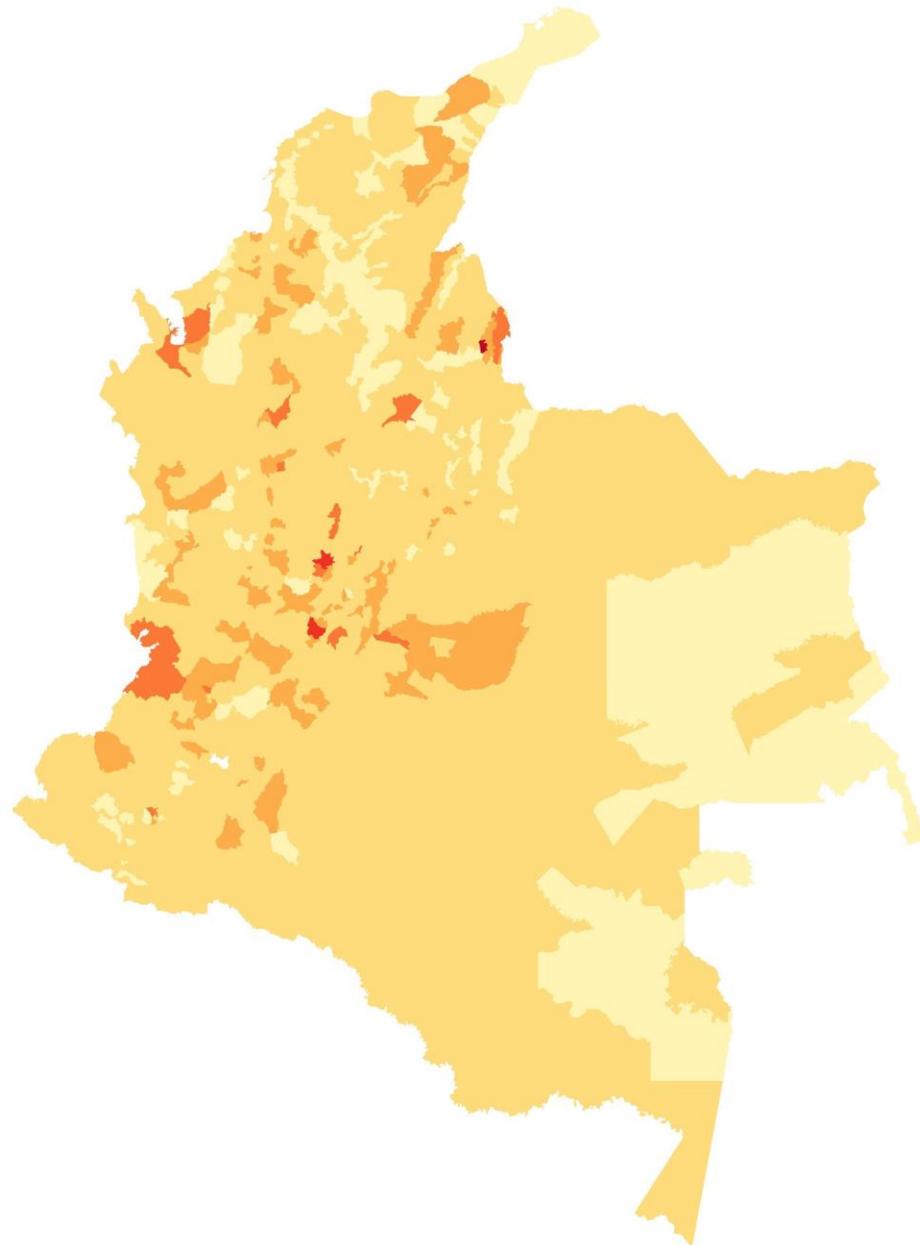
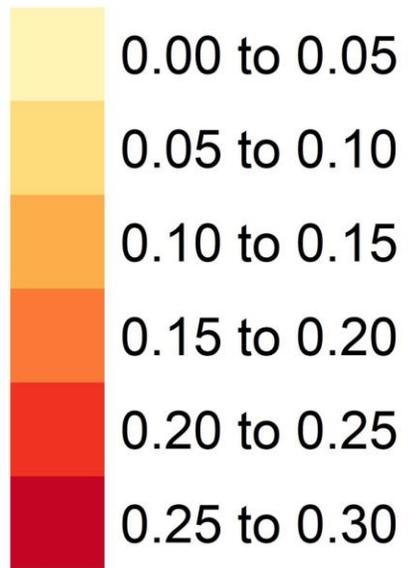
0.00 to 0.05
0.05 to 0.10
0.10 to 0.15
0.15 to 0.20
0.20 to 1.00

Maps for labor-force indicators

Unemployment Rate



Unemployment Rate

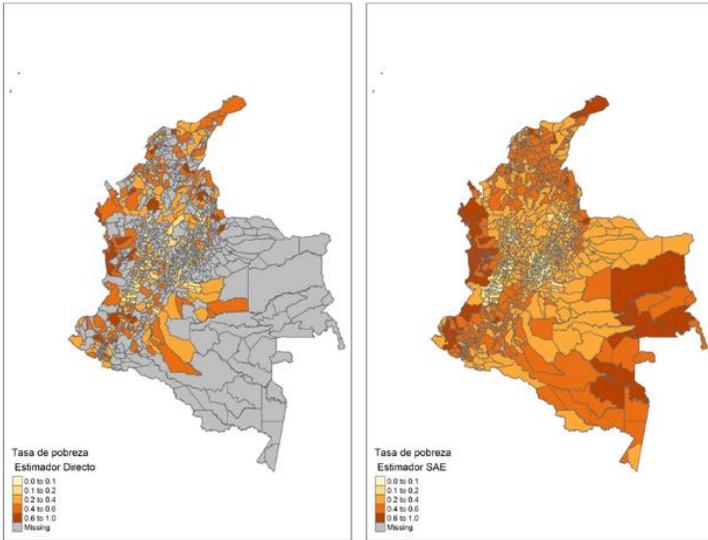


Un poco de historia

- En 2018, Isabel Molina escribió la primera nota de orientación sobre modelos de SAE con énfasis en los indicadores de pobreza de FGT.
 - Ese documento ahora se traduce al francés y al inglés y se utiliza en las capacitaciones de SAE en todo el mundo.
- Durante 2018 y 2019, la CEPAL inició una serie de capacitaciones presenciales en algunos países de ALC interesados.
 - Uruguay, El Salvador, Colombia, Chile.
 - Primeras versiones de scripts de R para modelos de nivel de área y de unidad.
- A partir de 2019, la CEPAL brindó apoyo técnico al UNFPA-LACRO en relación con el SAE de los indicadores de planificación familiar en algunos países de ALC.
 - Ahora la región cuenta con una metodología integral para estimar el uso de indicadores de PF y las necesidades insatisfechas de las mujeres fértiles.
 - Ecuador, Perú, Colombia, Guatemala, México and Jamaica.
 - El enfoque actual sigue el paradigma bayesiano.
- En 2020, la CEPAL capacitó a 18 países (+100 aprendices) de ALC en métodos de SAE.
 - La pandemia COVID19 profundizó la necesidad de contar con estadísticas más desglosadas.
 - No dejar a nadie atrás y centrarse primero en los más vulnerables.

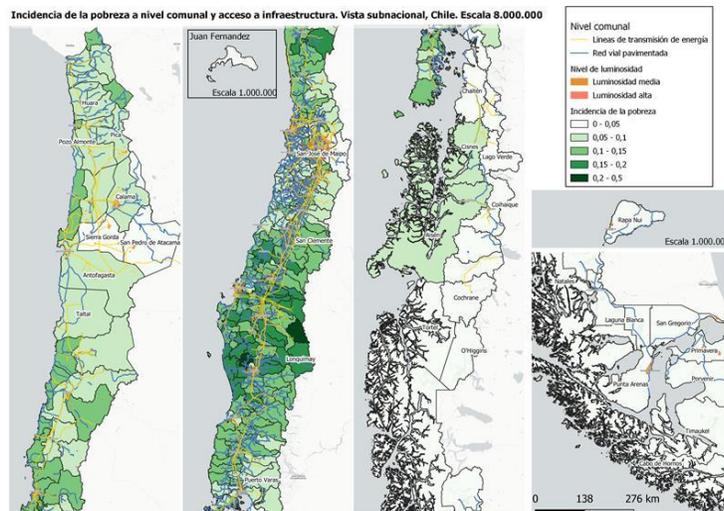


Un poco de historia

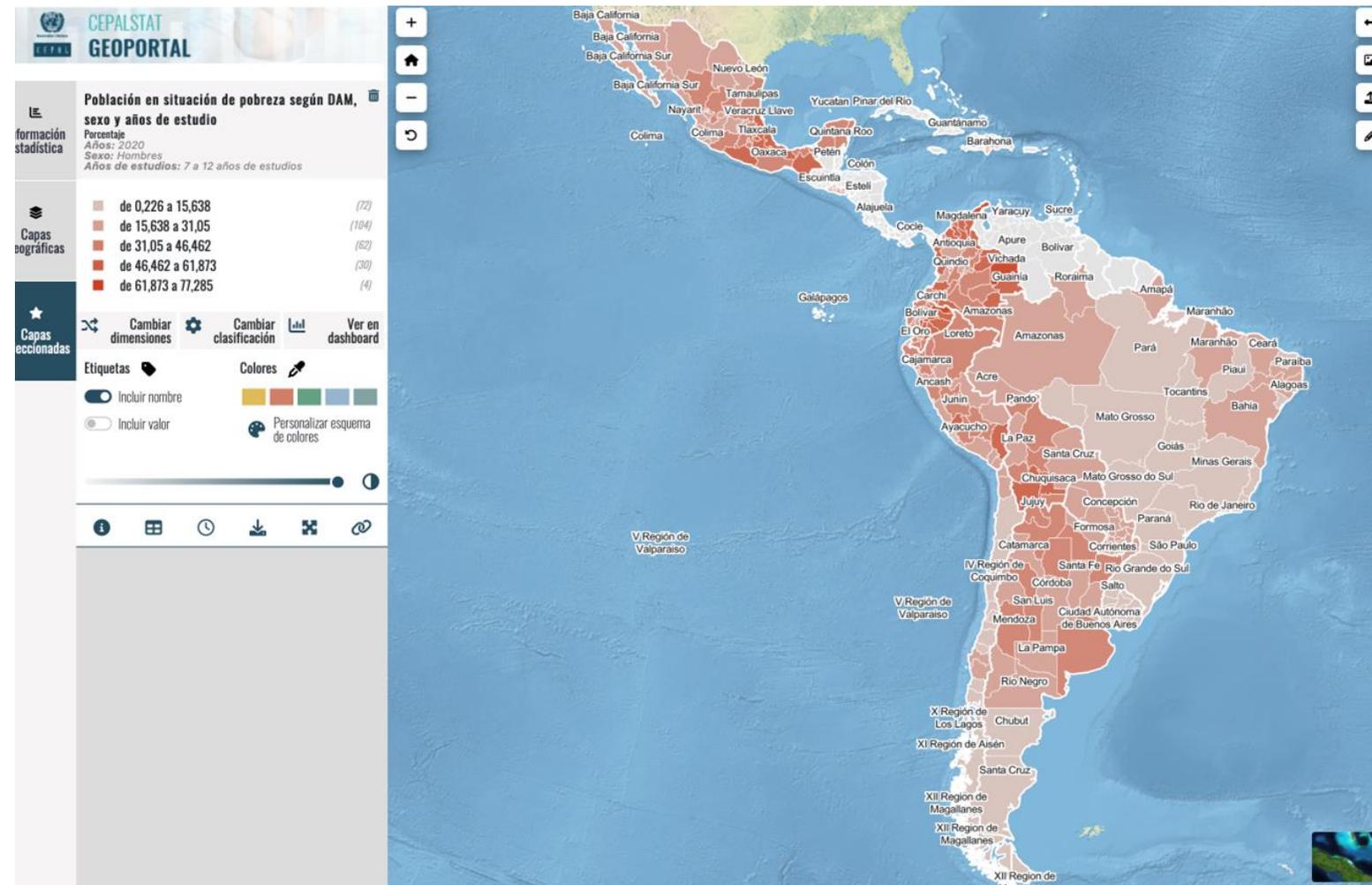


- Durante 2020 y 2021, la CEPAL apoyó proyectos específicos de mapeo de la pobreza de alto riesgo en algunos países de América Latina y el Caribe.

- Al gobierno de Chile por la revisión integral de la metodología de Mapeo de la Pobreza (modelos a nivel de área).
- El gobierno de Colombia también se benefició con una sólida metodología a nivel de unidad para el mapeo de la pobreza, que incluye scripts R para todo el proceso de SAE.



Un poco de historia



- En 2022, la CEPAL completó la automatización de la infraestructura estadística para la metodología bayesiana SAE para el mapeo de la pobreza en todos los países de América Latina.
 - Las estimaciones se actualizarán anualmente y se difundirán en el geoportal de la CEPAL.
- Desde 2022, la CEPAL se unió al DA13 y ha participado en el desarrollo de capacidades bayesianas de EAE para países beneficiarios específicos de América Latina y el Caribe.
 - Talleres presenciales en Ecuador, Guatemala, Perú, Chile, Colombia, Uruguay, República Dominicana, México y Brasil.

Un poco de historia

- En 2023, la CEPAL se unió a la FAO para apoyar al gobierno chileno en las estimaciones de FIES en municipios.
 - Las estimaciones FIES provienen de modelos de tipo Rasch.
 - Enfoques frecuentistas y bayesianos.
- En 2023, la CEPAL recibió una visita de la OIT y acordó explorar enfoques novedosos para la estimación de los métodos de SAE de mercado-trabajo en ALC.
 - Tasa de desempleo, tasa de participación, entre otros indicadores.
 - Modelos bayesianos con efectos aleatorios correlacionados.
- En 2023, la División de Estadística de las Naciones Unidas, la CEPAL y el UNFPA lanzaron la Capacitación Global sobre Métodos SAE con la participación de expertos internacionales.
 - Dos versiones de la Capacitación Global (a su propio ritmo y guiada).
 - Partha Lahiri, Isabel Molina, Pedro Silva, Mónica Pratesi, entre otros.
 - 10 módulos.

Gracias