

Seminario “Integración de fuentes de datos y estadísticas oficiales”

**Bloque: Integración de fuentes de
datos mediante modelos
estadísticos**

14 de Mayo 2024, Ciudad de México



Estimación de Áreas Pequeñas para Estadísticas Oficiales

Caso de la Estimación de la Escala FIES a nivel Municipal

Andrea Ramirez Pisco, Subdirectora del DANE

14 de Mayo 2024



Contexto actual de la producción estadística



Necesidad

- Formular política pública que atienda las necesidades de población y garantice el goce efectivo de derechos y que cuente con enfoque género, diferencial o interseccional.

Problema

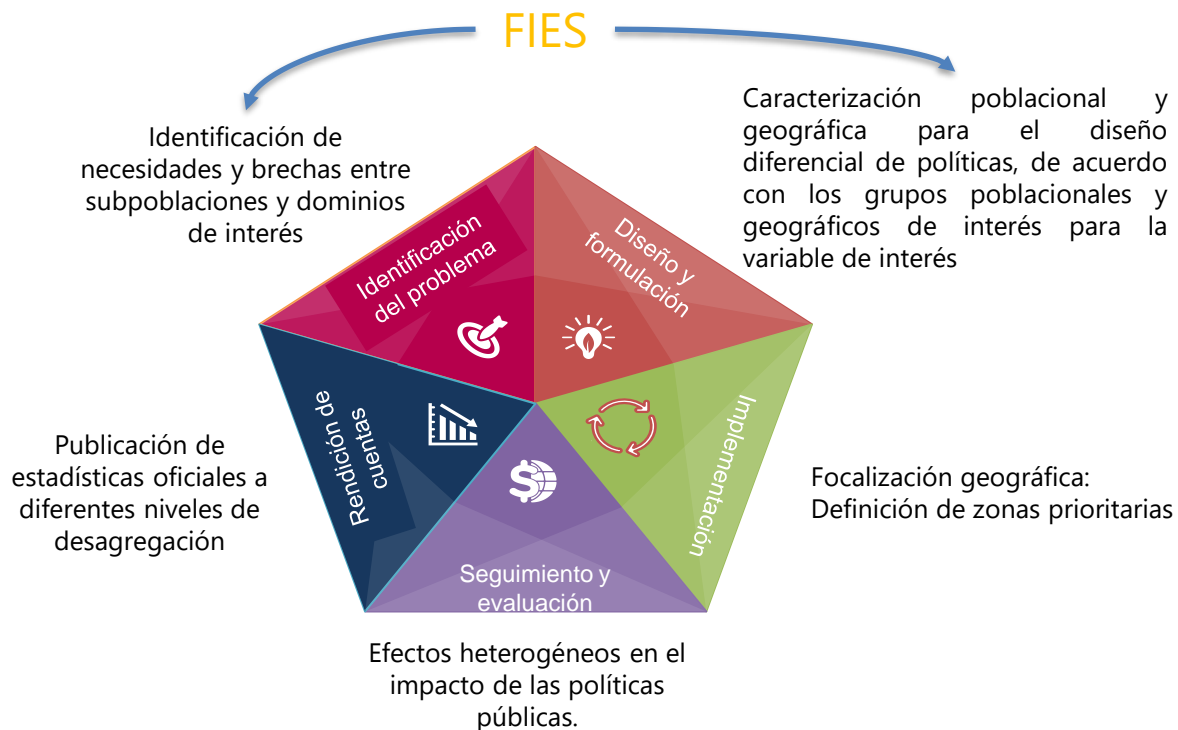
- La producción regular de estadísticas no permite el direccionamiento eficiente de las políticas al no tener el nivel de detalle suficiente para describir la población a niveles más detallados.
- Altos costos operativos para la producción de datos.
- Operativos complejos para la recopilación de información

Solución

- Implementación de metodologías estadísticas "novedosas" en la producción de estadísticas oficiales que permitan mayores niveles de desagregación de las variables de interés para la formulación de la política.

Formulación de política pública

En política pública se ha incrementado la demanda de información por parte entidades del estado, la academia, organismos no gubernamentales y la ciudadanía que demandan la **publicación de información estadística con mayor desagregación y detalle**. Esto con el fin de realizar ejercicios que permitan apoyar el ciclo de política pública, por ejemplo:



Niveles de desagregación definidos en el PND 2022-2026



De acuerdo con el Plan Nacional de Desarrollo 2022-2026 “**Colombia potencia mundial de la vida**”, se cuenta con los siguientes actores diferenciales para el cambio:



Género
Mujer
LGBTIQ+



Ciclo de vida
Niños y niñas
Jóvenes



Discapacidad



Víctimas



Grupos étnicos



Territorio
Municipios
Ruralidad
PDET



Campesinos



Economía Popular

Niveles de desagregación ODS



Los indicadores de los Objetivos de Desarrollo Sostenible deberían desglosarse, siempre que fuera pertinente, por:

- Ingreso
- Sexo
- Edad
- Raza
- Etnicidad
- Estado migratorio
- Discapacidad
- Ubicación geográfica

u otras características, de conformidad con los Principios Fundamentales de las Estadísticas Oficiales (resolución 68/261 de la Asamblea General).



Uso de SAE para la estimación de la prevalencia de inseguridad alimentaria (FIES) a nivel municipal

Escala de experiencia de inseguridad alimentaria (FIES)

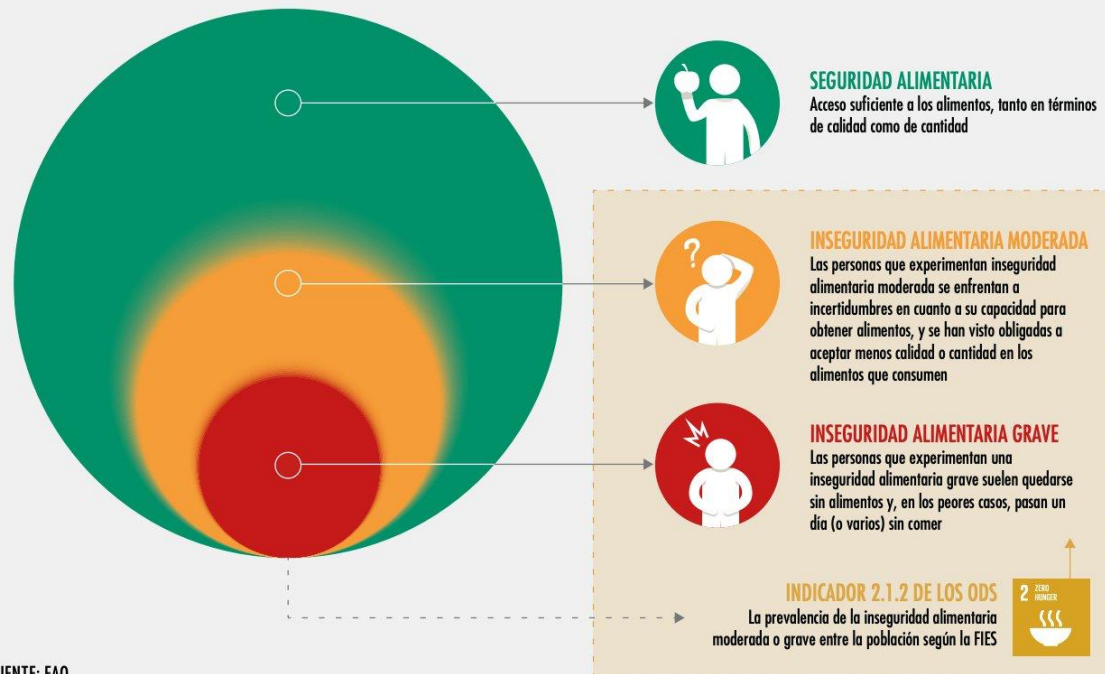
D



¿Qué es FIES?

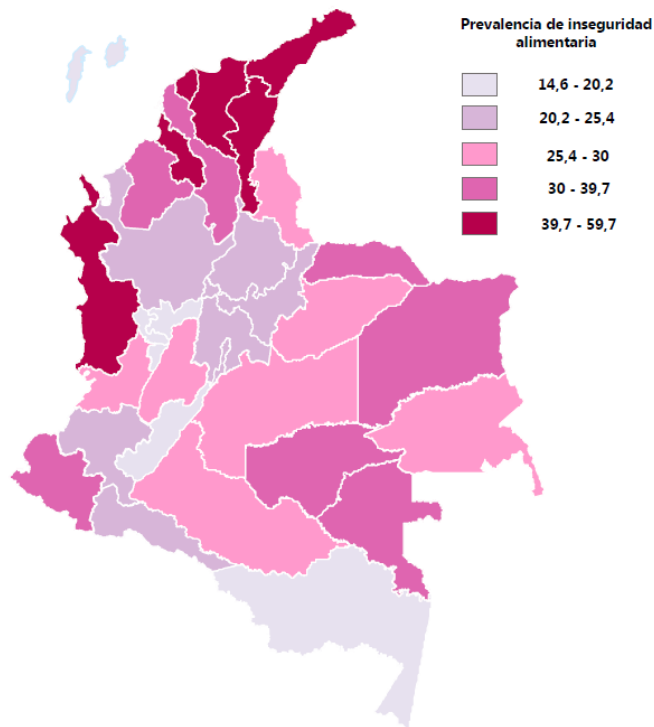
La Escala de Experiencia de Inseguridad Alimentaria (FIES, por sus siglas en inglés) es una herramienta de medición del acceso a los alimentos en términos de cantidad y calidad. A través de ocho preguntas la escala recoge la experiencia de la situación de inseguridad alimentaria de los hogares y personas, y proporciona una medición de la gravedad de la inseguridad alimentaria.

EXPLICACIÓN DE LOS NIVELES DE GRAVEDAD DE LA INSEGURIDAD ALIMENTARIA CUANTIFICADOS POR LA FIES EN EL INDICADOR 2.1.2 DE LOS ODS



FUENTE: FAO.

Resultados Iniciales de las Estimaciones FIES

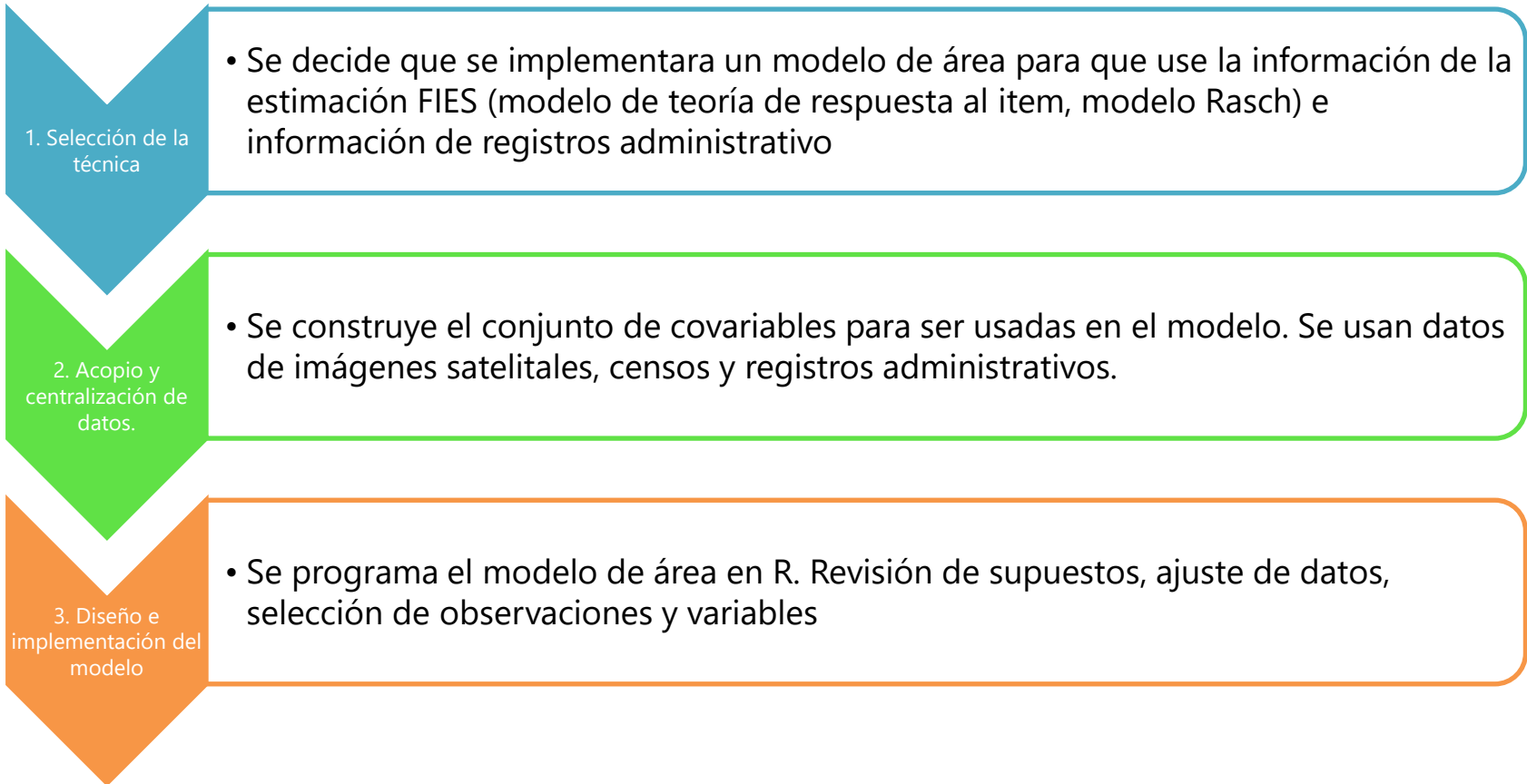


Con el enfoque muestral tradicional solo se lograron tener estimaciones de la prevalencia de la inseguridad alimentaria para 32 departamentos de Colombia

A partir de la publicación de FIES con la Encuesta Nacional de Calidad de Vida 2022 que permitió la estimación: Nacional, urbano/rural y departamental se identificó la necesidad de aumentar el nivel de desagregación geográfica de la variable con el fin de:

1. Mejorar el reporte del indicador 2.1.2 de ODS
2. Tener insumos para la focalización de la política de Hambre Cero liderado por la Vicepresidencia
3. Ser insumo para los Planes de Desarrollo Local

Proceso de Implementación de SAE



Integración de información

CNPV 2018

Estadísticas vitales 2021

SISBEN 2022

Censo Agropecuario 2014

EDUC 2022

EVAS

SISBEN 2022

BDUA 2022

MDM 2022

SIPSA

Se integraron diferentes fuentes de información a partir de censos y registros administrativos que permitieron construir 219 indicadores municipales asociados a las dimensiones de la seguridad alimentaria

Disponibilidad

Acceso

Consumo

Aprovechamiento y utilización

Calidad e inocuidad

+ algunas características de los municipios

Pasos implementados



01
FUNCIÓN GENERALIZADA DE VARIANZA
Se definió un nuevo modelo para el suavizamiento de la varianza usando $\text{Log}(\text{Var}) \sim B0+B1(n/N) +B2(n)$



02
FILTRO DE MUNICIPIOS
Aplicar reglas de filtros de municipios para asegurar la calidad de estimaciones directas. Se incluyó criterio para filtrar municipios con solo muestra rural o urbana



03
ANÁLISIS DE COVARIABLES
Eliminar variables duplicadas, de fuentes no confiables o provenientes de muestras, correlación entre variables mayor a 0,85 y VIF mayor a 10



04
STEPWISE
Usando el criterio de AIC para depurar variables se corre el algoritmo de StepWise como paso adicional se eliminan las variables no significativas en el modelo final



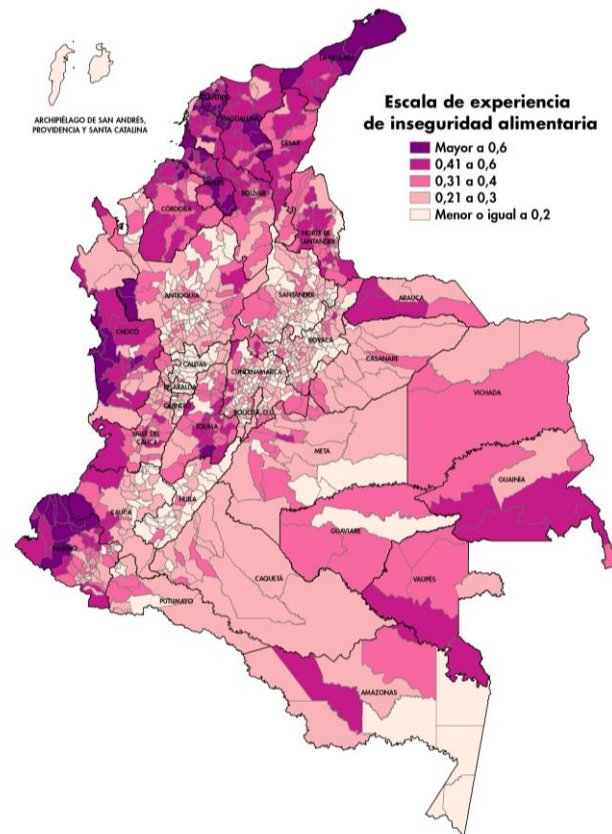
05
VALIDACIÓN
Se obtiene un modelo con $R^2=75\%$, normalidad en los errores y residuales. Además, se obtienen 1.048 municipios con CVLOG menor al 30%

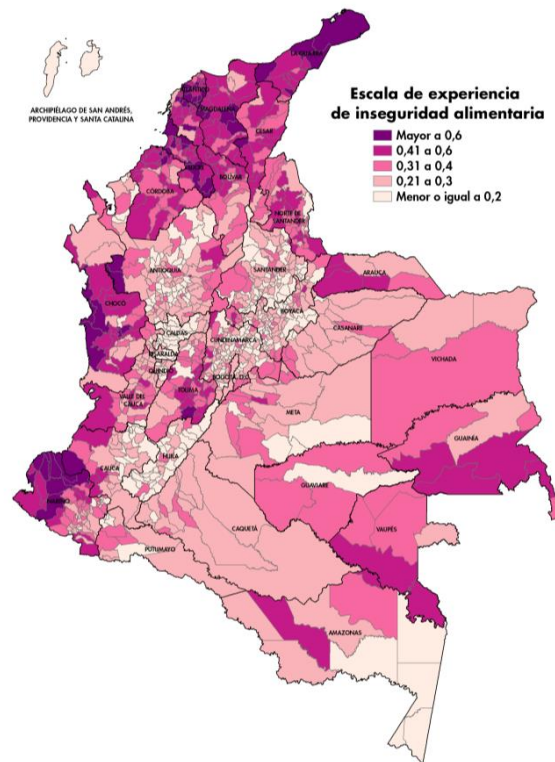
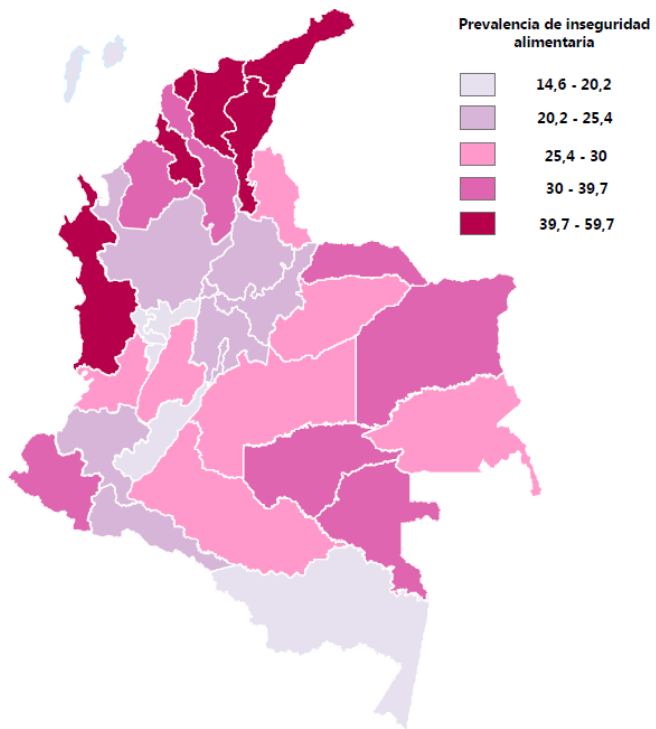
Estimaciones Municipales



Al aplicar el modelo de área (Fay – Herriot) para la estimación de la prevalencia de inseguridad alimentaria moderada-grave y haciendo uso de un total de 60 variables explicativas se obtuvo la estimación de 1.048 municipios con con CV menor al 30%.

Encontrando consistencia en las zonas esperadas con mayor incidencia del indicador el cuál fue validado por los expertos de FAO y CEPAL.





Algunas aplicaciones para la desagregación de variables en DANE



Mapa de pobreza multidimensional a nivel manzana/vereda

Integración de fuentes alternativas de información en el proceso estadístico



Actualmente el DANE mide:

- IPM a nivel departamento usando la encuesta de hogares anualmente
- IPM a nivel municipal usando el censo cada 10 años.

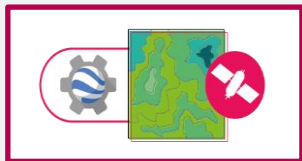
Indicador:

- IPM a nivel municipal anula

Sources:

- Encuesta de hogares
- Censo
- Datos Geoespaciales

Metodología:



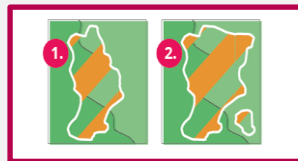
Recolectar

- Información geoespacial como : luces nocturnas, índice de vegetación, vías.



Entrada

- Clusters de la encuesta con la medida agregada de IPM



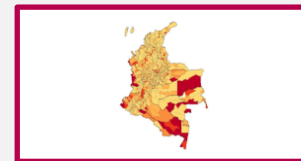
Modelo

- Modelo mixto generalizado (Geoestadístico)
- Modelo geoestadístico bayesiano



Estimación

- Población viviendo en pobreza a nivel de cluster



Resultados y validación

- Mapeo del IPM a nivel de cluster (manzanas/veredas)
- Rendimiento del modelo

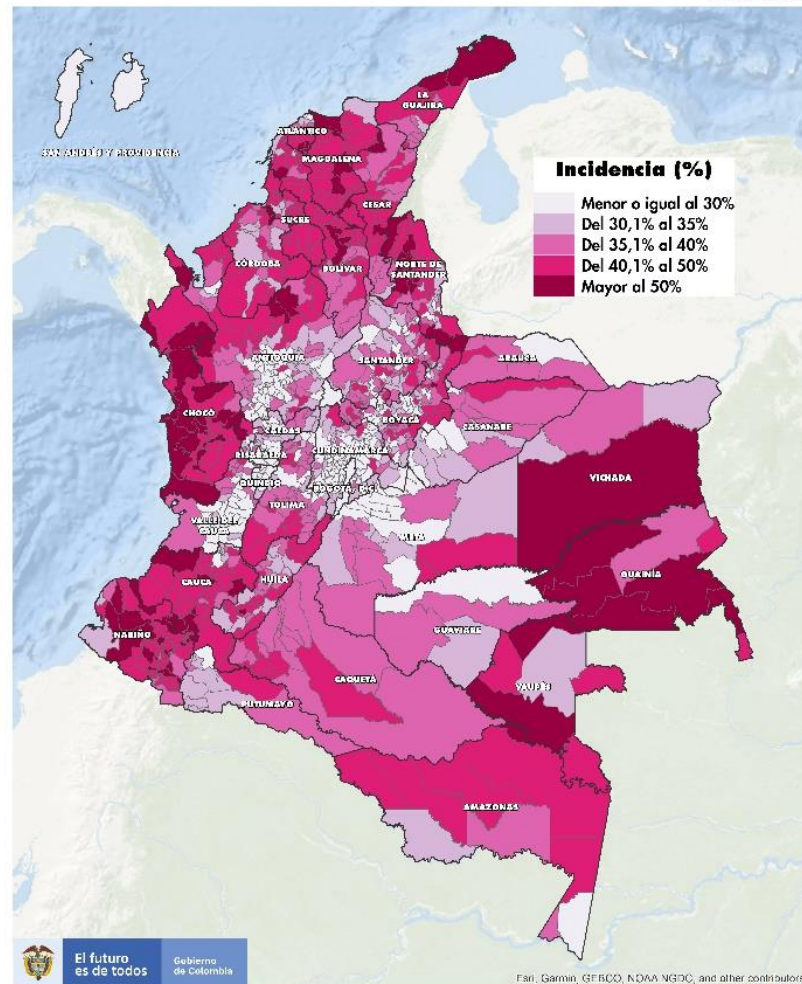
Mapa municipal de pobreza monetaria 2018

En cooperación con Data4Now se realizó la **estimación de pobreza monetaria para el año 2018 usando un modelo de área EBP (Empirical Best Predictor)**, usando el lenguaje de programación R. Para la construcción de covariables para el modelo se realizó la integración del Censo Nacional de Población y Vivienda 2018 junto con registros administrativos, como: SISBEN, BDUVA, Familias en Acción, Jóvenes en Acción, SIMAT, RELAB.

Como resultados, se encuentran que:

- Más allá de las usuales ciudades capitales de los departamentos, **el mapa muestra patrones de diferenciación geográfica entre los municipios en términos de la pobreza monetaria.**
- Estos patrones permiten distinguir aglomeraciones de municipios de baja o alta pobreza monetaria o, por el contrario, municipios que se encuentran en un extremo de la distribución rodeados por municipios en el extremo opuesto.

Este mapa no ha sido publicado dado que se encuentra en proceso de validación de los expertos de pobreza.



Estimación municipal de la emigración internacional en Colombia

En Colombia, los indicadores desagregados de alta calidad a nivel municipal solo están disponibles después de cada censo de población, es decir, cada 10 o 15 años. En los años intermedios, es normal que se produzcan cambios en los patrones de inmigración y en la intensidad del fenómeno. Por ello, las encuestas por muestreo proporcionan datos que permiten obtener estimaciones para períodos intermedios (estimaciones intercensales) y para períodos posteriores a los censos, en los que se dispone de los resultados de las encuestas (períodos poscensales).

Para la estimación la proporción de hogares con al menos un miembro habitual viviendo en el extranjero – PHMLA en períodos intercensales a nivel municipal se aplicó un modelo de área, Fay-Herriot, usando como fuente de información principal, la encuesta de demografía y salud del años 2015 (DHS2015).

Journal of Official Statistics, Vol. 37, No. 3, 2021, pp. 771–790. <http://dx.doi.org/10.2470/JOS-2021-0834>

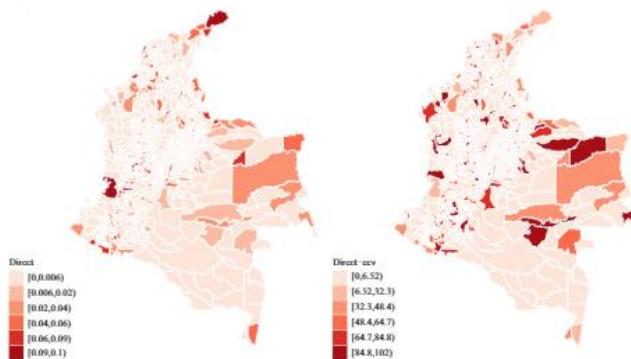
Fay-Herriot Model-Based Prediction Alternatives for Estimating Households with Emigrated Members

Jairo Fiquene-Patiño¹, César Cristancho², Mariana Ospina², and Domingo Morales Gonzalez²

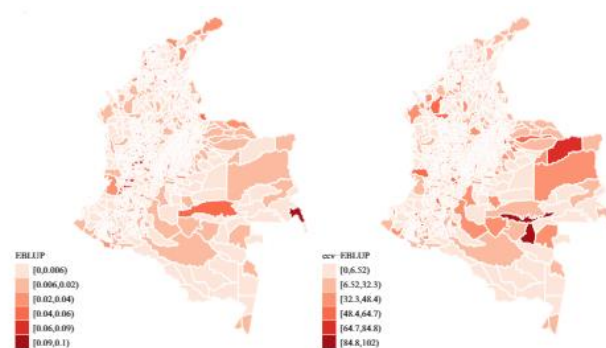
This article proposes a new methodology for estimating the proportions of households that had experience of international migration at the municipal level in Colombia. The Colombian National Statistical Office usually produces estimations of internal migration based on the results of population censuses, but there is a lack of disaggregated information about the main small areas of origin of the population that emigrates from Colombia. The proposed methodology uses frequentist and Bayesian approaches based on a Fay-Herriot model and is illustrated by one example with a dependent variable from the Demographic and Health Survey 2015 and covariables available from the population census 2005. The proposed alternative produces proportion estimates that are consistent with sample sizes and the main internal immigration trends in Colombia. Additionally, the estimated coefficients of variation are lower than 20% for municipalities for both frequentist and Bayesian approaches and large demographically relevant capital cities and therefore estimates may be considered to be reliable. Finally, we illustrate how the proposed alternative leads to important reductions of the estimated coefficients of variations for the areas with very small sample sizes.

Key words: Small area estimation; international migration; Fay-Herriot model; coefficient of variation; direct estimator; model-based estimator; hierarchical Bayes prediction.

Estimaciones directas y sus correspondientes CV



Estimaciones EBLUP y sus correspondientes CV



Uso de métodos anticonceptivos para población indígena



En cooperación con CEPAL y UNFPA se realizó la estimación del uso de métodos anticonceptivos de acuerdo con el grupo étnico desagregado a nivel departamental, para esto se utilizó un modelo de unidad, EBP (*Empirical Best Predictor*) y se usó como insumo el Censo Nacional de Población y Vivienda y la Encuesta de Demografía y Salud. Para lo cual se aplicaron los siguientes pasos:

01

Homologación de variables

02

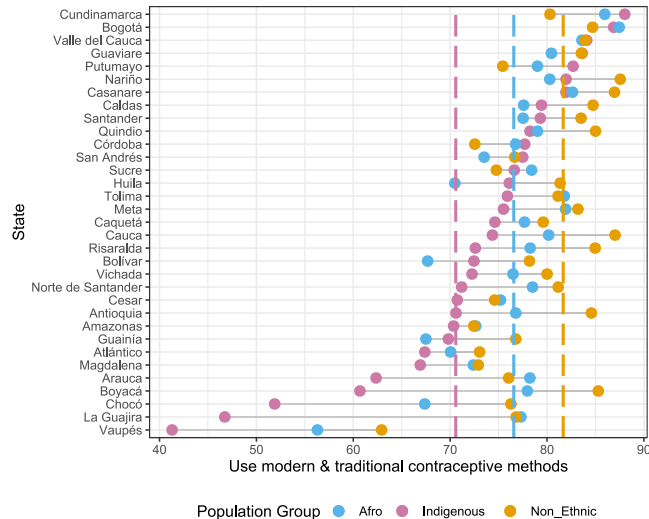
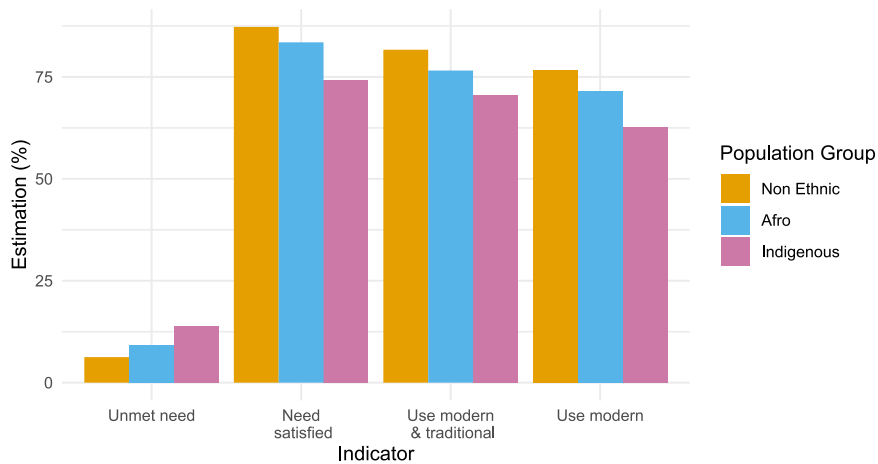
Usar los datos muestrales de la DHS para estimar el modelo

03

Usar los parámetros estimados con la información auxiliar del censo

04

Estimación de los dominios de interés



Gracias

Mayo 2024

