Leave No One Behind - International Organizations Efforts in Developing Countries

Instructions: Click on the link to access each author's presentation.

Chair: Andrés Gutiérrez

Participants:

Rolando Ocampo Alcántar: ECLAC experiences on disaggregating social indicators in LAC: from poverty to unemployment

Faryal Ahmed: Building national capacity on small area estimation

Mabely Diaz:* Dominican Republic Experiences to Disaggregation of Estimates based on Small Area Models

David Newhouse: The World Bank and Recent Small Area Estimation Endeavors

*Work presentation not available or non-existent

IAOS 2024 México

IPS title:

Leave No One Behind - International Organizations Efforts in Developing Countries

Participants:

Mr Rolando Ocampo Alcántar - ECLAC experiences on disaggregating SDG indicators: from poverty to unemployment

With over 35 years of statistical expertise, Rolando Ocampo has dedicated his career to promoting the use of statistical and geospatial information for decision-making and public policy. He holds degrees from top universities and has played significant roles in Mexico's National Institute of Statistics and Geography (INEGI) and international statistical organizations. As of August 2019, he serves as the Director of the Statistics Division at the Economic Commission for Latin America and the Caribbean (ECLAC).

Ms Faryal Ahmed – Leave No One Behind - International Organizations Efforts in Developing Countries

Faryal Ahmed is a statistician and coordinator for the Data For Now initiative, that leverages innovative data sources, methodologies, technologies to fill urgent data gaps to respond to national policy priorities, at the United Nations Statistics Division. She supports countries in strengthening their national statistical systems to fill data gaps across various domains, including environment, urban planning, food security and agriculture. Faryal also contributes to the work on citizen data and was part of the core team to organize the United Nations World Data Forum in 2020 and 2021.

Mr David Newhouse – Modernizing small area estimation

David Newhouse is a distinguished Senior Economist at the World Bank. Holding a PhD in Economics from Cornell University, he has made significant contributions to the field through multiple publications focused on poverty and the use of small area estimation. With a notable career trajectory, David has held various key positions within the World Bank, including Senior Economist, Labor Economist, and Senior Economist Consultant.

Welcome to the sesion:

Leave No One Behind - International Organizations Efforts In Developing Countries

To achieve the "Leave No One Behind" imperative, crucial for implementing the Sustainable Development Goals (SDGs) and the 2030 Agenda, National Statistical Offices must overcome challenges through innovative data integration methods, combining data from diverse sources like surveys, censuses, records, imagery, and social media. The objective is not only disaggregating official statistics geographically but also tailoring data for scattered vulnerable populations. In this session, International Organizations (IOs), such as the United Nations, ECLAC, and The World Bank, showcase their essential role in aiding developing nations to enhance their data systems. They provide support in technical assistance, capacity-building, data harmonization, and financial resources, contributing to higher-quality data for evidence-based policymaking and sustainable development. The session delves into real-world experiences of IOs supporting countries in adopting a data integration culture for SDG indicators, highlighting the challenges, particularly in small area estimation capacity-building.

Mr Rolando Ocampo Alcántar - ECLAC experiences on disaggregating SDG indicators: from poverty to unemployment

With over 35 years of statistical expertise, Rolando Ocampo has dedicated his career to promoting the use of statistical and geospatial information for decision-making and public policy. He holds degrees from top universities and has played significant roles in Mexico's National Institute of Statistics and Geography (INEGI) and international statistical organizations. As of August 2019, he serves as the Director of the Statistics Division at the Economic Commission for Latin America and the Caribbean (ECLAC).

The ECLAC presentation focuses on disaggregating SDG indicators to ensure no one is left behind. It highlights the importance of breaking down data by income, sex, age, race, ethnicity, migration status, disability, and location. Due to limitations in survey sample sizes, ECLAC employs Small Area Estimation (SAE) models, integrating data from censuses, administrative records, and satellite imagery. The process involves six stages: standardizing databases, updating intercensal counts, defining models, predicting indicators, validating models, and generating maps. Data sources include household surveys from BADEHOG, census data from the Population Division, and remote sensing data from platforms like Google Earth Engine. ECLAC ensures reliability through strict validation and benchmarking, creating detailed social indicators for 17 Latin American countries to support targeted policies and sustainable development efforts.

Ms Faryal Ahmed – Leave No One Behind - International Organizations Efforts in Developing Countries

Faryal Ahmed is a statistician and coordinator for the Data For Now initiative, that leverages innovative data sources, methodologies, technologies to fill urgent data gaps to respond to national policy priorities, at the United Nations Statistics Division. She supports countries in strengthening their national statistical systems to fill data gaps across various domains, including environment, urban planning, food security and agriculture. Faryal also contributes to the work on citizen data and was part of the core team to organize the United Nations World Data Forum in 2020 and 2021.

The presentation "Leave No One Behind - International Organizations Efforts In Developing Countries" by Faryal Ahmed focuses on building national capacity for Small Area Estimation (SAE) to improve data for Sustainable Development Goals (SDGs). The Inter-Secretariat Working Group on Household Surveys (ISWGHS) and the Inter-agency and Expert Group on Sustainable Development Goal Indicators (IAEG-SDGs) are key in this effort, developing tools and methodologies for data disaggregation. The SAE4SDG Toolkit offers practical guidance, case studies, and software resources to help countries implement SAE for official statistics. Challenges include limited resources, technical capacity, and quality input data. The presentation highlights the need for a supportive environment, policy-driven needs, and continuous capacity building, exemplified by successful applications in countries like

Colombia, Chile, Indonesia, and Jamaica. Future steps include enhancing training, utilizing geospatial data, and fostering collaboration for more effective SAE implementation.

Mr David Newhouse – Modernizing small area estimation

David Newhouse is a distinguished Senior Economist at the World Bank. Holding a PhD in Economics from Cornell University, he has made significant contributions to the field through multiple publications focused on poverty and the use of small area estimation. With a notable career trajectory, David has held various key positions within the World Bank, including Senior Economist, Labor Economist, and Senior Economist Consultant.

David Newhouse's presentation "Modernizing Small Area Estimation" highlights advancements in combining survey data with auxiliary data to enhance statistical power and improve data quality. Modernization includes using geospatial data from satellites, providing updated, granular, and unbiased indicators, and exploring social media and call detail records. Improved SAE methods address heterogeneity, selection bias, and diagnostics.

Beyond poverty mapping, SAE enhances accuracy in labor, health, education, and population outcomes, though some indicators remain hard to estimate accurately. Tools like the Geolink R package and Povmap aid in merging geospatial data with survey data and offer advanced features for SAE applications. The presentation underscores rapid advancements in geospatial data, the need for ongoing methodological improvements, and the importance of accessible tools and documentation for broad adoption and capacity-building.

IAOS-ISI 2024
MEXICO CONFERENCE

UNITED NATIONS
ECLAC

ECLAC experiences on disaggregating SDG indicators: from poverty to unemployment

Rolando Ocampo
Director
Statistics Division
ECLAC

INEGI
isi International Statistical Institute
IAOS IMPROVING OFFICIAL STATISTICS

Why do we do the things we do?


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

SDG focuses on dissagregation

Target 1.2. By 2030, reduce at least by half the proportion of men, women and children of all ages living in poverty in all its dimensions according to national definitions.

- Indicator 1.2.2: Proportion of men, women and children of all ages living in poverty in all its dimensions according to national definitions.

Target 8.5. By 2030, achieve full and productive employment and decent work for all women and men, including for young people and persons with disabilities, and equal pay for work of equal value.

- Indicator 8.5.1: Average hourly earnings of employees, by sex, age, occupation and persons with disabilities

LEAVE NO ONE BEHIND

Fundamental principle of data disaggregation

Sustainable Development Goal indicators ***<u>should be disaggregated</u>***, where relevant, by income, sex, age, race, ethnicity, migration status, disability and geographic location, or other characteristics, in accordance with the Fundamental Principles of Official Statistics.

General Assembly resolution - 68/261

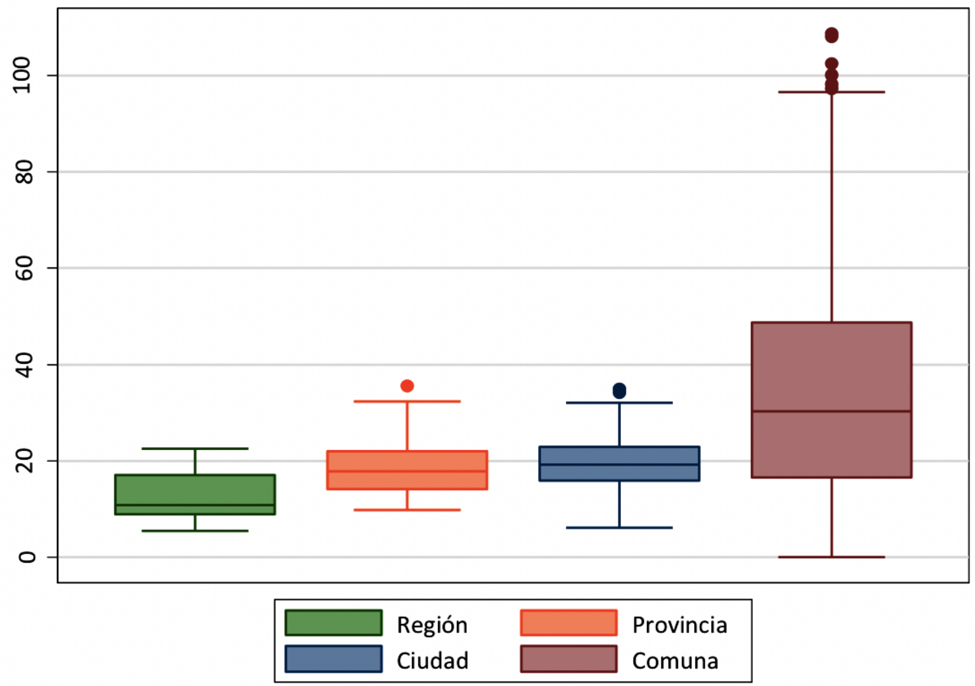
IOS-ISI 2024 MEXICO CONFERENCE

SAE models in ECLAC

Surveys are not enough

Surveys depend on large sample size and a proper sampling strategy (sampling design and estimator).

They also rely on a robust inferential system that provides precise and exact estimation in planned domains.



The problem

When the sample size does not allow obtaining reliable direct estimates for some domains of interest, the following options can be addressed:

1. Increase the sample size: this option raises costs, and it is unfeasible.
2. Use statistical methodologies that involve external auxiliary information to obtain reliable estimates (not direct) in the subgroups of interest while keeping the survey sample size.

Types of small areas

It is necessary to resort to external auxiliary information (censuses, administrative records, satellite imagery) so that together (surveys and external data) a precise inferential system can be built.

- Geographical units could include states, provinces, departments, or municipalities.
- Specific subgroups might involve combinations such as age × sex × ethnicity × immigration status.

UN-ECLAC uses SAE models to integrate data from different sources.

ECLAC approach to SAE models (1)

The implementation of the ECLAC methodology for poverty maps in each of the 17 countries followed these six stages:

- Stage 1. Standardization and homologation of covariates in the databases (censuses and household surveys). Adaptation of satellite imagery as state-level covariates.
- Stage 2. Updating intercensal counts related to covariates preserving the census structures while updating marginals from the household survey.
- Stage 3. Definition of the models for indicators, considering possible interactions, selection of auxiliary variables and estimation of model coefficients.

ECLAC approach to SAE models (2)

The implementation of the ECLAC methodology for poverty maps in each of the 17 countries followed these six stages:

- Stage 4. Prediction of indicators on censal poststrata and small areas, estimation of the MSE based on MCMC replicas.
- Stage 5: Validation of model assumptions and benchmarking using ECLAC direct estimates at the national, urban, and rural levels.
- Stage 6: Generation of maps for 17 countries of Latin America.

Model for poverty

Molina & Rao (2010) estimated poverty rates and gaps at the crossroads of gender and province in Spain using the following model.

Link model:

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

Conditional estimation:

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

Model for labor force indicators

Molina, Saei and Lombardía (2007) proposed a methodology based on the assumption that the sample totals individuals follow a multinomial logit model with random area effects.

They considered a multinomial vector that counts the number of sampled unemployed (y1), employed (y2), and inactive (y3) individuals within each small area group.

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

Integrating data for SAE purposes

Different sources of data

1. Household survey data is obtained from ECLAC's Household Survey Data Bank (BADEHOG):
   - Maintained by the Statistics Division on an annual basis since the 90s. This repository contains household surveys from 18 Latin American countries.

2. Census data is accessed through ECLAC's census data bank:
   - Maintained by the Population Division (CELADE). The Population Division has an ongoing activity of disseminating census information derived from its data bank.

3. Remote sensing data is obtained through Google Earth Engine and Open Street Maps.

Country	Survey	Year
ARG	Permanent Household Survey (EPH)	2019
BOL	National Household Survey	2020
BRA	National Survey by Continuous Household Sample	2020
CHL	National Socioeconomic Characterization Survey (CASEN)	2020
COL	Large Integrated Household Survey	2020
CRI	National Household Survey (ENAHO)	2020
DOM	National Continuous Labour Force Survey (ENCFT)	2020
ECU	National Survey on Employment, Unemployment and Underemployment (ENEMDU)	2020
GTM	National Survey on Living Conditions	2014
HND	Multipurpose Household Survey	2019
MEX	National Household Income and Expenditure Survey (ENIGH)	2020
NIC	National Household Survey on Living Standard Measurement	2014
PAN	Multipurpose Survey	2019
PER	National Household Survey - Living Conditions and Poverty	2020
PRY	Continuous Permanent Household Survey (EPHC)	2020
SLV	Multipurpose Household Survey	2020
URY	Continuous Household Survey	2020

Country	Census	Year
ARG	National Census of Population, Households, and Housing	2010
BOL	Population and Housing Census	2012
BRA	Demographic Census	2010
CHL	Population and Housing Census	2017
COL	National Census of Population and Housing (CNPV)	2018
CRI	X National Census of Population and Housing	2011
DOM	IX National Census of Population and Housing	2010
ECU	VII Population and Housing Census	2011
GTM	XII National Census of Population and VII Housing Census	2018
HND	XVII Population and Housing Census	2013
MEX	Population and Housing Census	2020
NIC	Population and Housing Census of Nicaragua	2005
PAN	Census 2010	2010
PER	XII Census of Population, VII Housing Census, and III Indigenous Communities Census	2017
PRY	II National Indigenous Census of Population and Housing	2002
SLV	VI Population Census and V Housing Census	2007
URY	Population Census	2011

Satellite imagery

Among the main advantages of information based on remote sensing is the ease of access to data with deep geographic coverage that is impossible to obtain by traditional means such as surveys or administrative records.

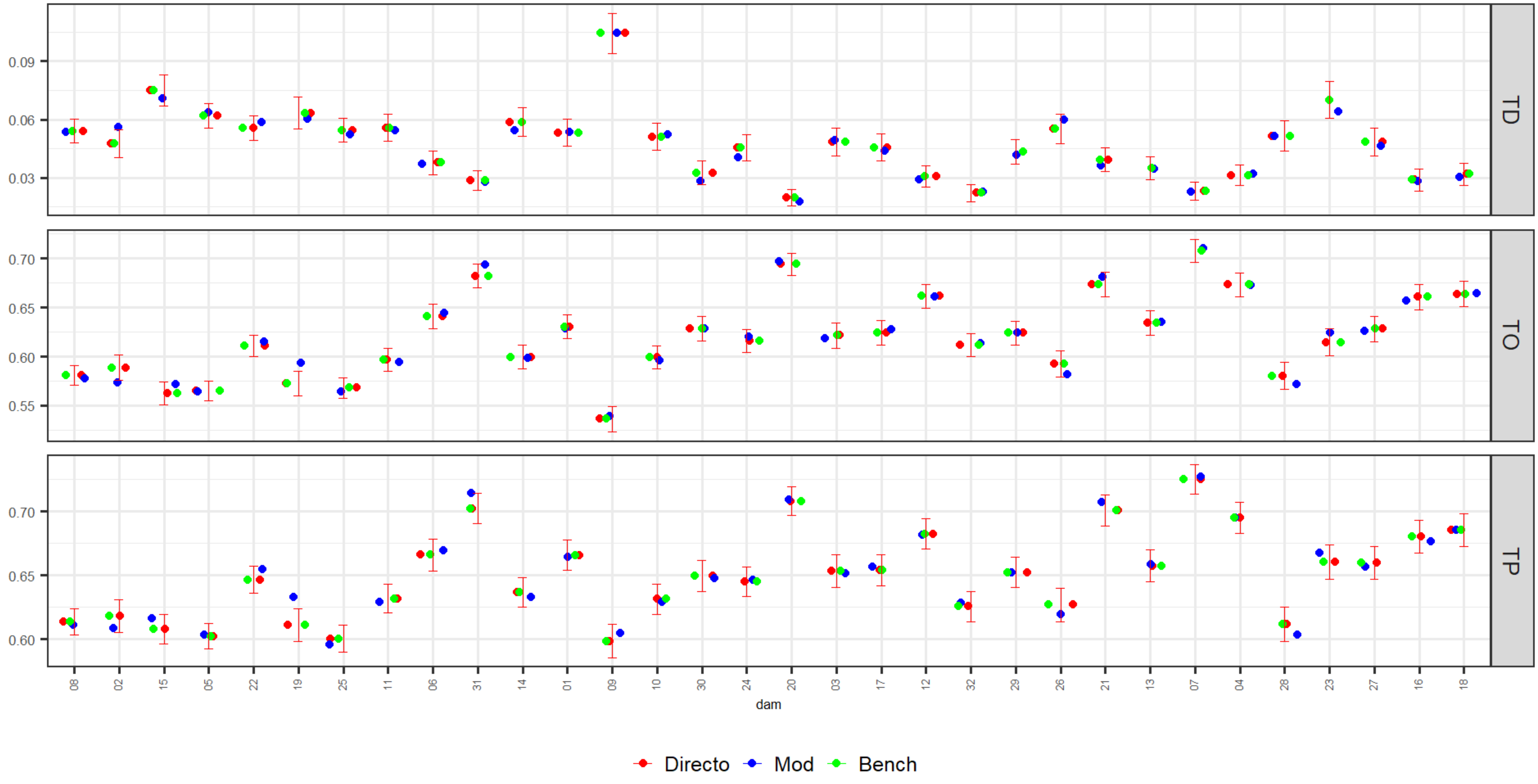
- We access this information trough Google Earth Engine, which provides facilities to analyze and obtain this data through the Javascript and Python programming languages, and recently in R with the *rgee* package.

- The Statistics Division of ECLAC is currently developing the Image Bank of Products from Satellite Image Processing Project, aimed at processing remote sensing data to be used in a variety of projects, including SAE estimates of poverty and unemployment.

Benchmarking and model asssessment

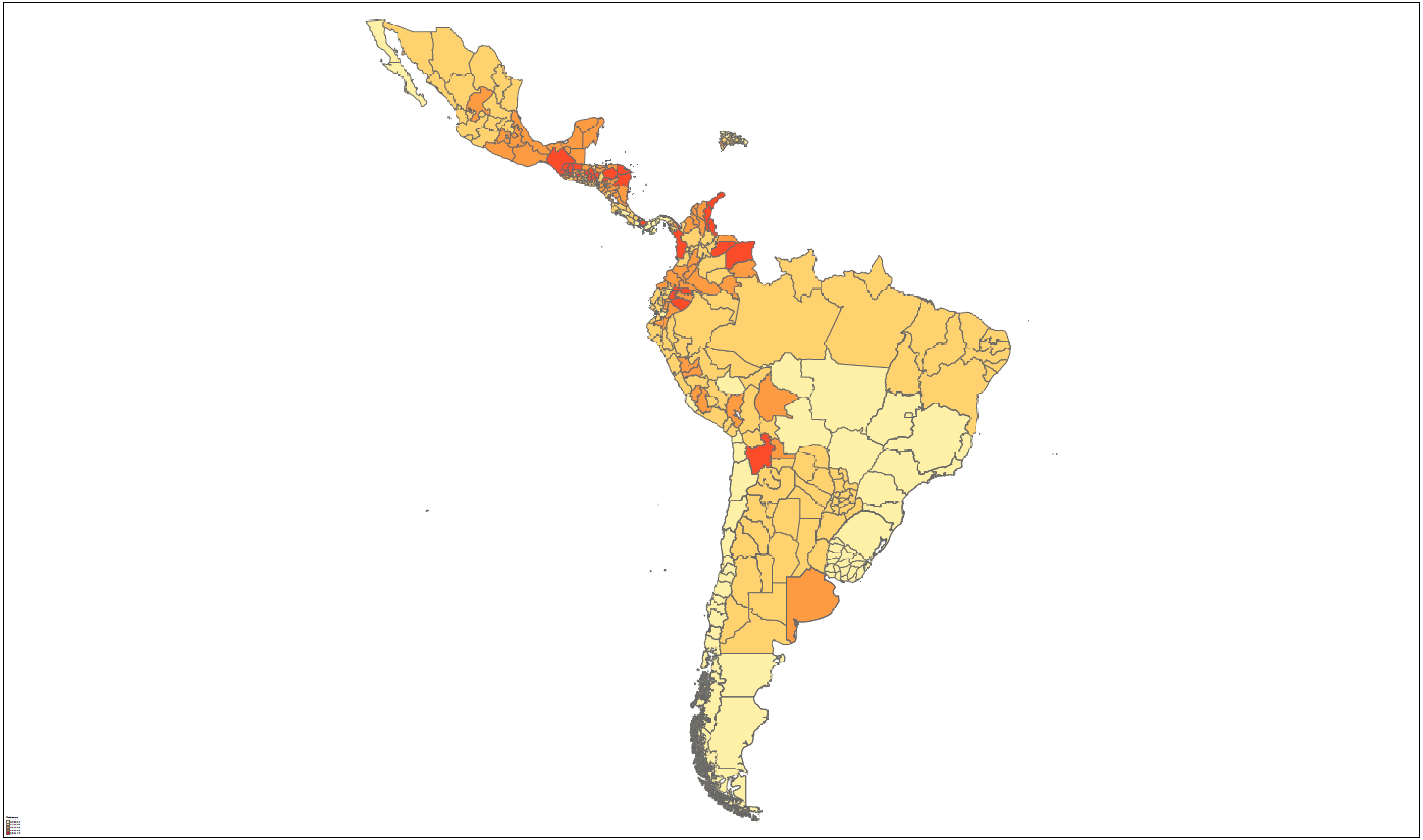
Once the chains have reached convergence, we use the direct estimates at the national and regional levels to benchmark the small area estimates. This process is carried out for each iteration of a Markov Chain monte Carlo process.

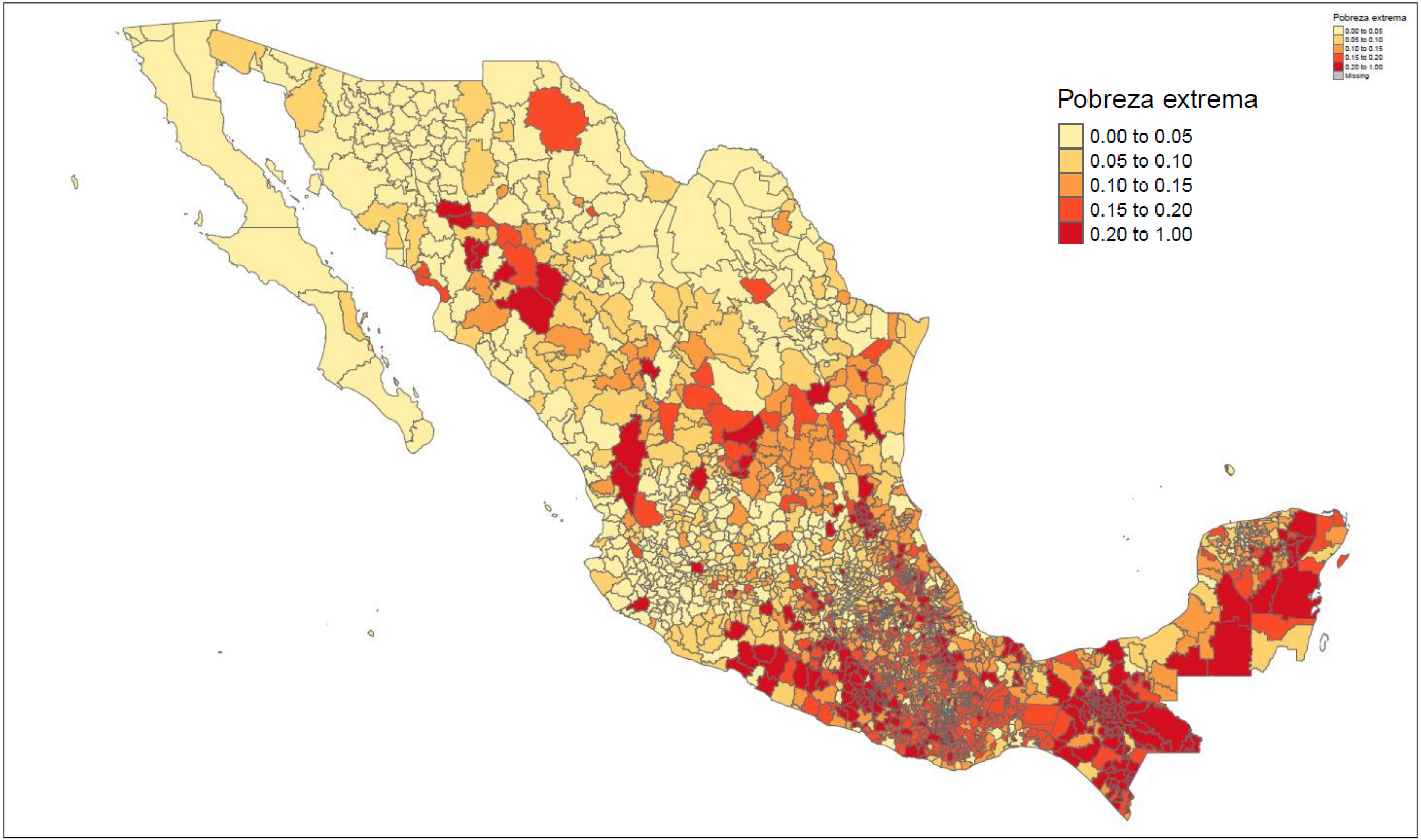
We carry out a strict validation and model assessment process to detect possible violations of the model assumptions. Tests used assess for normality and heteroskedasticity. Also, we use Cook’s distances as measures of the influence of the observations in the sample.



TD
TO
TP
0.09
0.06
0.03
0.70
0.65
0.60
0.55
08 02 15 05 22 19 25 11 06 31 14 01 09 10 30 24 20 03 17 12 32 29 26 21 13 07 04 28 23 27 16 18
dam
Directo
Mod
Bench

Maps for poverty indicators

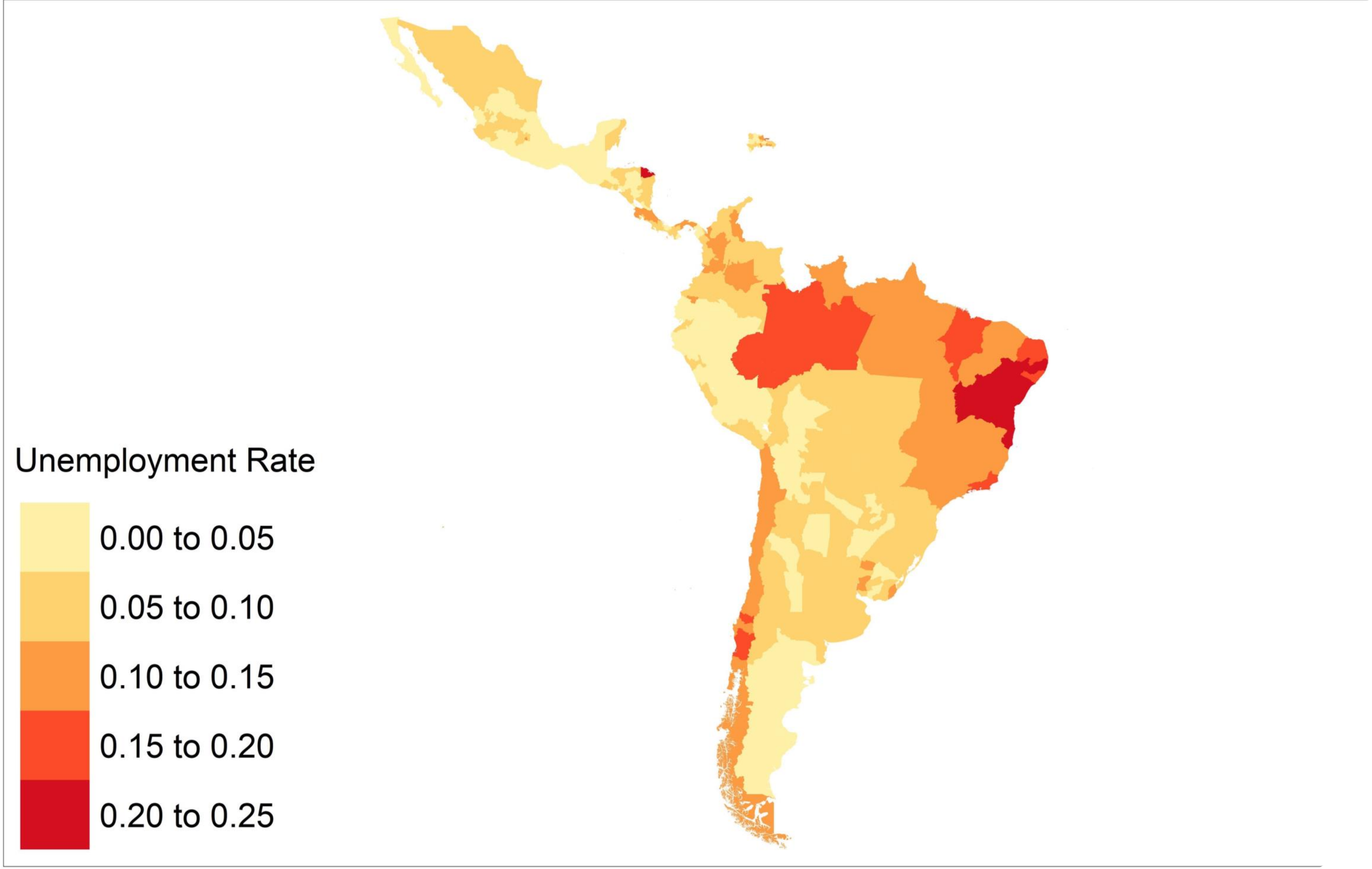


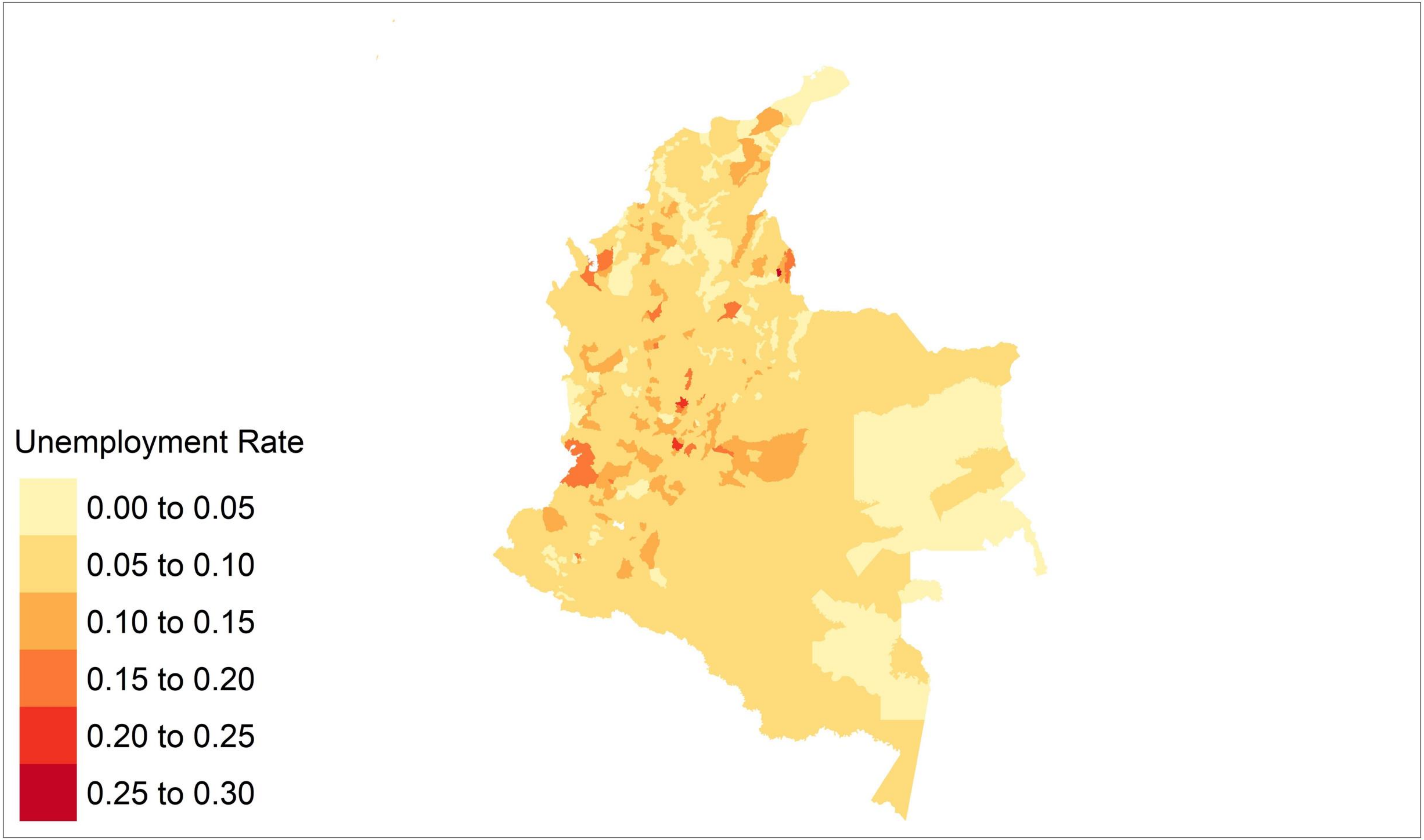


Pobreza extrema
0.00 to 0.05
0.05 to 0.10
0.10 to 0.15
0.15 to 0.20
0.20 to 1.00
Missing
Pobreza extrema
0.00 to 0.05
0.05 to 0.10
0.10 to 0.15
0.15 to 0.20
0.20 to 1.00

Maps for labor-force indicators

IOS-ISI 2024
MEXICO CONFERENCE


Unemployment Rate
0.00 to 0.05
0.05 to 0.10
0.10 to 0.15
0.15 to 0.20
0.20 to 0.25


Unemployment Rate
0.00 to 0.05
0.05 to 0.10
0.10 to 0.15
0.15 to 0.20
0.20 to 0.25
0.25 to 0.30

In summary…

Our efforts

- UN-ECLAC has made several efforts to provide the region with estimates of social indicators at granular levels for all countries in Latin America.

- We use different sources of data to integrate them into SAE models (for poverty and the labor force).

- It is a huge effort that will bridge the gap in information needed not only in geographical areas but also in subgroups of interest (ethnicity, age, education, disability, etc.).

- We are complying with the Leave No One Behind mandate of the 2030 agenda.

Thank you

IAOS-ISI 2024
MEXICO CONFERENCE

Leave No One Behind - International Organizations Efforts In Developing Countries

INEGI

International Statistical Institute

IAOS
IMPROVING OFFICIAL STATISTICS

Building national capacity on small area estimation

Faryal Ahmed, Statistician and Project Coordinator
United Nations Statistics Division

unstats.un.org/iswghs

Outline

- ❑Inter-Secretariat Working Group on Household Surveys (ISWGHS) & IAEG-SDGs
- ❑Toolkit on using SAE for SDG indicators
- ❑Capacity building activities on SAE: challenges and opportunities
- ❑Next steps

The ISWGHS: a primer

- ❑ Established in 2015 under the aegis of the UNSC
- ❑ Objectives:
  - ❑ Improve coordination of household surveys
  - ❑ Advance cross-cutting survey methodology
  - ❑ Enhance communication and advocacy
- ❑ Governance
  - Membership: 11 international agencies + 10 (rotating) member states
  - Secretariat: UN Statistics Division
  - Current co-chairs: WB and UNW
- ❑ Work through time-bound Task Forces, led by and with contribution from members and non-member experts.

ISWGHS

Inter-agency and Expert Group on Sustainable Development Goal Indicators (IAEG-SDGs)

The 2030 Agenda for Sustainable Development

- ❑ A global blueprint for people, planet, prosperity , peace and partnerships, now and in the future
- ❑ 17 Goals, 169 targets and "Leaving no one behind" principle

The IAEG-SDGs :

- ❑ Composed of 28 Member States (and representatives of regional commissions, regional and international agencies and CSOs are observers)
- ❑ Developed the global indicator framework for SDGs (**231 indicators)**

IAEG-SDGs workstream on data disaggregation:

- ❑ Compilation of existing guidelines and methodologies on data disaggregation
- ❑ Preparation of Handbook on data disaggregation for SDGs
- ❑ Task Force on Small Area Estimation (joint with ISWGHS)

ISWGHS

Positioning household survey for the next decade

Organized around **8 technical priorities**:

1. Enhancing the interoperability and integration of household surveys
2. Designing and implementing more inclusive, respondent-centric surveys
3. Improving sampling efficiency and coverage
4. Scaling up the use of objective measurement technologies
5. Building capacity for CAPI, phone, web, and mixed-mode surveys
6. Systematizing the collection, storage, and use of paradata and metadata
7. Incorporating machine learning and artificial intelligence for data quality control and analysis
8. Improving data access, discoverability, and dissemination.

Plus:

Foster stronger **enabling environment**:
at national and global level

https://content.iospress.com/articles/statistical-journal-of-the-iaos/sji220042

ISWGHS

The SAE4SDG Toolkit

- **The Toolkit on Using Small Area Estimation for SDGs** (https://unstats.un.org/wiki/display/SAE4SDG/) in Wiki is a space to provide information on methods to produce disaggregated data through small area estimation.

- **Goal:** To provide practical tools with accompanying case studies for countries to use SAE for SDG monitoring.

- **Objectives:**
  - Using SAE methods to improve SDG data availability for vulnerable population groups
  - Offering practical guidance and country case studies
  - Guiding on the enabling environment for using SAE for official data production
  - Providing a space for partners to document and disseminate their SAE methodologies

ISWGHS

What the SAE Toolkit Offer

- Many countries have experimented with SAE in the past but few were able to transform from experiment to official production. The Toolkit:
  - Finds out why this is happening?
  - Establishes a close link of SAE to SDG monitoring
  - Provides hands-on exercise, including "semi-synthetic" data (national data + noises) and programming guide.
  - Incorporates national examples and case studies through two angles: (a) documenting the lessons learnt and challenges of countries in using SAE for official data production; and (b) illustrating SAE practices for indicators under different SDG goals.
  - Includes main challenges and enabling environment to move from SAE experiment to official production, based on our discussion with national statistical offices.
  - Provides an up-to-date and comprehensive list of SAE software packages in major languages (R/Stata/SAS/Python).

ISWGHS

Guiding through steps with practical examples

8.5.2 Unemployment rate

R Code

- User needs
- Data availability
- Specification
- Analysis & Adaptation

Evaluation & Benchmarking

To evaluate the domain indicators, the model is fitted and the MSE and the CV as measure for the uncertainty of the estimates are estimated. The estimation of the MSE and CV is triggered by setting the parameter MSE to "TRUE". For the transformed area-level model with bias-corrected backtransformation, a bootstrap MSE is provided. The parameter B controls the number of bootstrap iterations. It is advisable to set B to a minimum value of 100 in order to obtain reliable MSE estimates.

Precision, accuracy and reliability › Expand source

The estimated regional indicators (the unemployment rate in this example) with its MSE and CV can be obtained in the form of a table. Generally, the CV should be used with caution when the indicator of interest is a ratio since really low point estimates can also be the reason for large CVs. In these cases, it is recommendable to focus on the MSE.

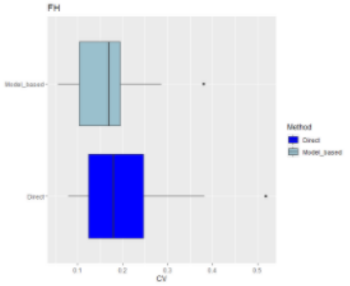
In this example, it can be seen that the CV of the model-based estimate (FH) is generally lower than for the direct estimate. However, there are also cases where the CV is slightly larger. One reason could be that the number of bootstrap iterations is too low.

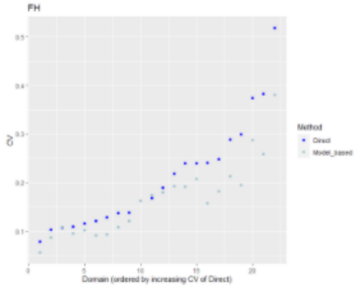
MSE and CV per domain › Expand source

The model-based estimates are commonly compared with the results of direct estimates. The function compare_plot in emdi provides some plots for this comparison.

Comparison with direct estimation › Expand source

Comparing direct with model-based estimates helps to evaluate if the model-based estimates are more reliable than the direct estimates measured in terms of the MSE or the CV. The boxplots confirm that the model-based estimates have lower CVs overall. Approximately, 75% of the model-based domain estimates show a CV below 20%. It is also apparent that the increase in efficiency is not huge. Furthermore, the second plot shows that there are also domains where the CV of the model-based estimates is larger than the one of the direct counterpart.





When comparing the direct and model-based point estimates, it can be seen that these do not differ strongly from each other.

Case studies covering different SDG goals/indicators

Goal 1. End poverty in all its forms everywhere

› Case studies

Goal 2. End hunger, achieve food security and improved nutrition and promote sustainable agriculture

› Case studies

Goal 3. Ensure healthy lives and promote well-being for all at all ages

› Case studies

Goal 4. Ensure inclusive and equitable quality education and promote lifelong learning opportunities for all

› Case studies

Goal 5. Achieve gender equality and empower all women and girls

› Case studies

ISWGHS

SAE methodologies used by countries and international agencies

Dashboard / SAE4SDG 38 views

SAE practices

Created by Haoyi Chen, last modified on May 04, 2021

Asian Development Bank

FAO

UNICEF

US Census Bureau

Asian Development Bank

Created by Haoyi Chen, last modified by Arturo Jr M. Martinez on May 04, 2021

Brief introduction of the organisation

ADB is committed to achieving a prosperous, inclusive, resilient, and sustainable Asia and the Pacific, while sustaining its efforts to eradicate extreme poverty. Established in 1966, it is owned by 68

A description of the SAE work within the organisation

In 2017, the Asian Development Bank (ADB) launched the Data for Development project which aims to support the statistical capacity of national statistics offices (NSOs) in Asia and the Pacific, help
to monitor the Sustainable Development Goals (SDGs). This component focuses on strengthening the capacity of NSOs to generate fine-grained data for policies and programs targeted to vulnera

One of the outputs of this component is a guide on disaggregation of official statistics, which includes an inventory of various small area estimation (SAE) methodologies to yield granular data for
explains SAE techniques with examples of how the easily accessible R analytical platform can be used to implement them, particularly to estimate indicators on poverty, employment, and health ou

Reference:

- Asian Development Bank. Introduction to Small Area Estimation Techniques: A Practical Guide for National Statistics Offices

Future work on SAE

The guide compiles various SAE techniques and worked examples on how to implement the methodology, which were covered in a series of country training workshops provided to the staff of sev
disaggregated data requirements of the SDGs. Furthermore, since its publication in May 2020, several researchers and academics have reported the usefulness of the guide in their work.

Moving forward, the team will continue exploring potential areas of collaboration with national statistical systems who may need technical assistance in building capacity on the application of SAE

Like Be the first to like this

Write a comment...

US Census Bureau

Created by Haoyi Chen, last modified on May 04, 2021

Introduction

One of the most famous programmes on small area estimation for official statistics is the Small Area Income and Poverty Estimates (SAIPE) Program led by the US Census Bur
discussion with the SAIPE team at the US Census Bureau as well as other reference materials.

How to motivate SAE - how did you convince the government to use small area estimates?

Answer: Prior to SAIPE, all local level income and poverty information can only be produced from the decennial census long-form. This means that small area estimates on po
based largely on "the number of children aged 5 to 17, inclusive, from families below the poverty level on the basis of the most recent satisfactory data, ..., available from the
the Department of Commerce, unless the Secretaries of Education and Commerce determine that the use of updated population data would be "inappropriate or unreliable."

From the description above, three distinct features stand out:

1. A legal act is in place that requires that the Secretary of Education distribute Federal funds based on data produced at county and school district level, unless data are "
2. The legal act also specifies that such data should be produced by the Department of Commerce that houses the US Census Bureau
3. Funding of an external expert panel to provide quality check

Therefore this is really a "top-down" approach where the law requires that quality data are to be used for policymaking, distributing Federal fund in this case. The program is

Input data

Surveys that provide poverty data: Current Population Survey (CPS) through 2004 and American Community Survey starting in 2005.

Administrative data:

- US Federal income tax data
- Supplemental Nutrition Assistance Program (SNAP) participants data
- Supplemental Security Income (SSI) recipiency rate

Data from the Census Bureau Population Estimates Program are used to construct denominators of several of the regression covariates.

Source: An Overview of the US Census Bureau's Small Area Income and Poverty Estimates (SAIPE Program), Bell, Basel and Maples, 2015

Input data quality reflection

Quality of the input data is important. One administrative data that was considered but not used is the Free and Reduced-Price Lunch Data. Studies showed such data are no

One reflection is on how household surveys could be better designed to allow good small area estimation. For example, CPS sample that collected poverty data are relatively

Adjustment made on the model and estimates

Improvements of small area estimates are made over time, by refining models and incorporating new or updated data sources. Since its inception SAIPE program has made m

ISWGHS

Challenges in using SAE for official statistics

- Lack of interest and support from the top management
- Lack of dedicated resources for SAE research and implementation
- Lack of in-house technical capacity
- Lack of proper input data (access to/poor quality of admin data source)
- Reluctance about the use of model-based estimates (vs. survey estimates that are design-based/model-assisted)
- Difficulties in communicating the technical aspects to users

ISWGHS

Challenges in using SAE for official statistics (cont.)

- *"We did an experiment using small area estimation method for poverty but the results were not consistent with our own estimates so we did not pursue it again."*
- *"We do not have good input data source for SAE - census data are outdated, and administrative data sources do not have good coverage and lack proper auxiliary variables."*
- *"SAE method is complicated and we are not comfortable with independently developing the method."*
- *"It is very difficult to convince the managers to use model-based estimates."*
- *"Producing SAE requires a lengthy period of looking for input data, finding the right auxiliary variables, testing different models and their assumptions and validating the estimates."*

Source: UNSD conversations with NSOs

ISWGHS

Enabling environment for SAE

- *Establishing a clear and focused objective that links SAE to data use for policymaking*
- *Building the legal foundation for using SAE for official data production*
- *Fostering an environment for research and development*
- *Design-based versus model-based estimates: a changing culture in the national statistical offices*
- *Input data for SAE*
- *Maintaining a high and fit-for-purpose quality standard*
- *Collaboration*
- *Capacity building*
- *Transparency in releasing methodology and communicating quality*

ISWGHS

Lessons learnt: driven by needs for key policies and funding decisions

- *Colombian National Development Plan 2018-22 made it mandatory to redesign the national monetary transfer programs* (Jóvenes en Acción and Familias en Acción), for population in poverty and in extreme poverty. This needs poverty data at municipal level. (Colombia)

- *In 2009, the law of the Fondo Común Municipal (FCM) required the Ministry to provide poverty rate estimates every 2 years for all comunas in the country. Funding to all comunas will be allocated based on such data.* (Chile)

- *The 2005-2009 BPS Strategic Plan for Statistical Development defined "the development of an efficient and low-cost methodology, which allows for the creation of small area and local specific statistics data" as one of the main activities to support government decentralization (*Indonesia*)*

- *The Cabinet of the Government of Jamaica made a request for the Statistical Institute of Jamaica to use small-area estimation for poverty mapping, to produce poverty data for smaller geographical areas within the country. (*Jamaica*)*

- Improving America's Schools Act: *"the number of children aged 5 to 17, inclusive, from families below the poverty level on the basis of the most recent satisfactory data, ..., available from the Department of Commerce"* (US)

ISWGHS

Lessons learnt: access to good quality input data

- ❑ Access to auxiliary data sources (e.g., administrative data), regularly
- ❑ Input data are of good quality:
  - Coverage, accuracy and timeliness
  - Availability of auxiliary variables that have good prediction power for the outcome indicator

Table 20.5 Initial set of auxiliary variables reviewed for their possible inclusion as comuna level auxiliary variables in the area level model.

Name of the auxiliary variable	Institution responsible for data collection	Frequency of publication of the data
1. Subsidio Familiar	Unidad de Prestaciones Monetarias, Ministerio de Desarrollo Social	Monthly and yearly
2. Subsidio al Pago del Consumo de Agua Potable y Servicio de Alcantarillado de Aguas Servidas	Unidad de Prestaciones Monetarias, Ministerio de Desarrollo Social	Monthly and yearly
3. Bono Chile Solidario	Unidad de Prestaciones Monetarias, Ministerio de Desarrollo Social	Monthly and yearly
4. Subsidio de Discapacidad Mental	Unidad de Prestaciones Monetarias, Ministerio de Desarrollo Social	Monthly and yearly
5. Pensión Básica Solidaria (vejez e invalidez)	Unidad de Prestaciones Monetarias, Ministerio de Desarrollo Social	December
6. Aporte Previsional Solidario (vejez e invalidez)	Unidad de Prestaciones Monetarias, Ministerio de Desarrollo Social	December
7. Bonificación al Ingreso Ético	Unidad de Prestaciones Monetarias,	Monthly and yearly

Source: Example from Chile, Casas-Cordero, Encina and Lahiri (2016)

ISWGHS

Capacity building on SAE

- A joint effort of ECLAC-UNSD-UNFPA: https://learning.officialstatistics.org/user/index.php?id=103
  - Reading materials
  - Recorded videos (50 videos with about 10-15 minutes for each video), organized in 10 modules
  - Evaluation materials including weekly computer-graded assessments, two mid-term projects, and a final project
  - R program language code that can be used for SAE modelling
- Opened in August 2023
  - Self-paced students on the platform: 460
  - Guided learning sessions with an extra 1.5-hour per week to provide guidance: 200 students registered and we are currently supporting around 120 students from Asia, Africa and Latin America (with ECLAC, ESCAP and ECA)

ISWGHS

Offering more and better training

- High demand: continuing the eLearning course guided training 2024:
  - SIAP will be offering one session for Asia and the Pacific
  - One for English-speaking African countries and one for Latin America and the Caribbean
  - French translation soon to be available, for Francophone African countries (self-paced)
- Reflecting on the learning experiences: R skills, linear model, busy schedules, sometimes the interested students do not really work on the area, course material very intense
- Improving the training experiences:
  - Reducing the complexity of the project assignments, to cater to different levels of students
  - Doing more intensive follow-ups/reminders with students on homework assignments/video watching
  - Making certain modules elective for more advanced students
  - Preparing Syllabus that has specific grading/marking requirements
  - Extending the course completion period by 1-2 more weeks to allow extra time for projects

ISWGHS

Geospatial data for SAE: a review of its potential, limitations and effectiveness

1. An overview of SAE method, why and the audience of the review
2. Input data: geospatial data and training data
3. Geospatial SAE methods
4. Skills and tools to apply the methods
5. Future research and work

- ❑ A draft available: here; will finalise end 2024
- ❑ Partners: World Bank, SAE expert, IAEG-SDGs, GGIM-ISGI

ISWGHS

Geospatial data for SAE: hands-on guidance

- To develop a step-by-step guidance on:
    1. Accessing geospatial data for SAE
    2. Selecting the types of data to use
    3. Illustrating with datasets
- Regional training: Asia/Pacific and Africa

Thank you

MODERNIZING SMALL AREA ESTIMATION

WORLD BANK GROUP

David Newhouse
Development Economics Data Group

August 22, 2024

Modernization agenda

- Small Area Estimation (SAE) enables surveys to “borrow strength” from more comprehensive auxiliary data
  - Combines interpretability of surveys with greater statistical power of auxiliary data
  - Useful for evaluation, targeting, and assessing survey data quality

- Long history of SAE at World Bank
  - Elbers, Lanjouw, and Lanjouw (2003) proposed combining survey and census data to “borrow strength” from census data when estimating poverty
  - WB recently produced updated guidelines on poverty mapping (Corral et al, 2023)
    - Main change is to encourage use of EBP method proposed by Molina and Rao (2010)
  - EBP based on huge statistics literature on empirical bayes / mixed effects models
    - Carter and Rolph, 1974, Efron and Morris 1979, Fay and Herriot 1979, Laird and Ware 1982, Battese Harter and Fuller 1988, Jiang and Lahiri, 2006, Pinheiro and Bates 2006 among others
  - EBP predictions combine direct survey estimates and synthetic model predictions in an efficient way.

Modernization agenda

- Most WB SAE applications still utilize survey and census data
    - Estimates updated when new census data becomes available every ten years or so

- Efforts to modernize
    1. New sources of auxiliary data (“big data”)
    2. New and improved statistical methods
    3. More indicators beyond poverty mapping
    4. New tools to harness recent innovations

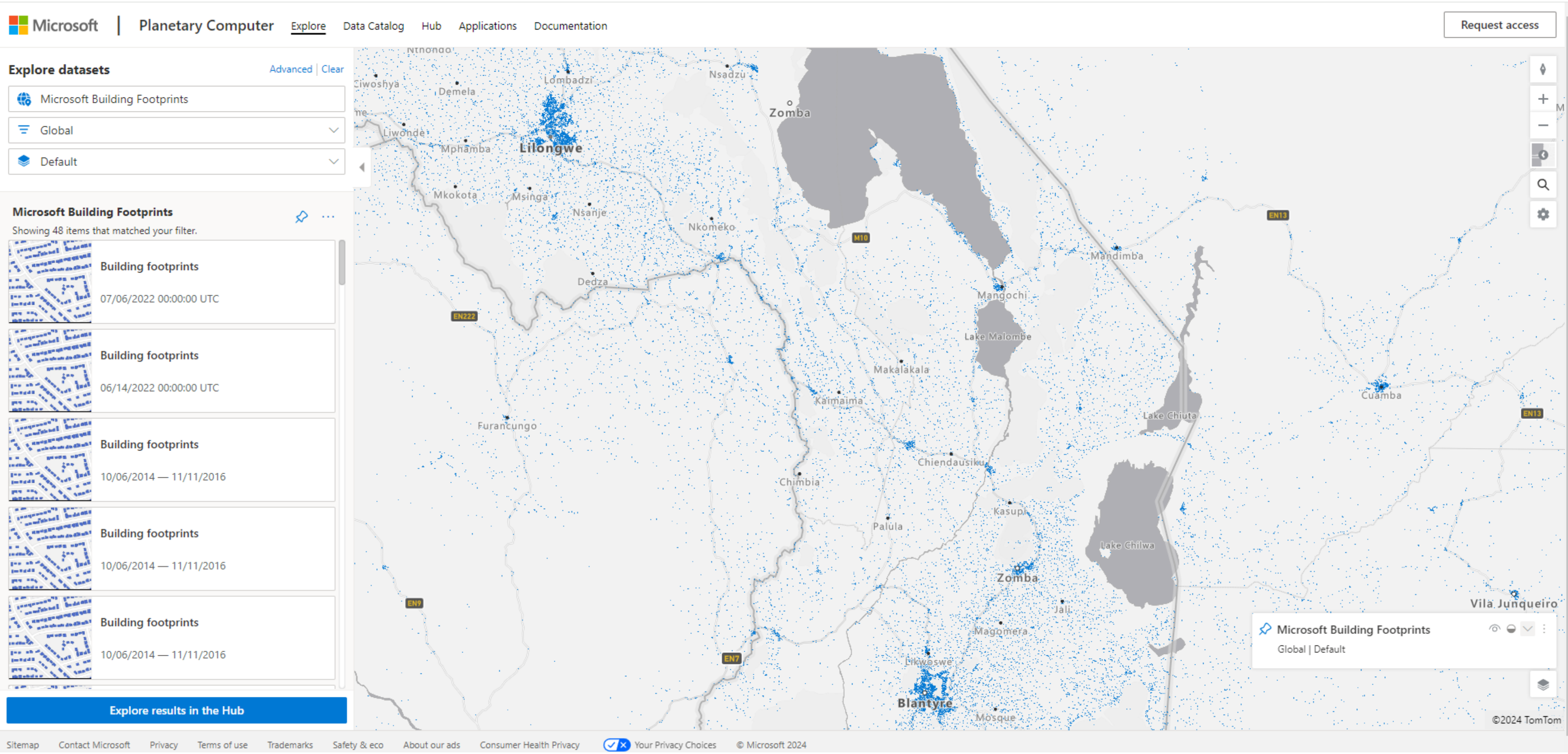
- Great interest in applying these innovations to generate useful and trustworthy data
    - Many applications: SAE, impact evaluation, policy evaluation, bias correction, improved statistical estimates, etc.
    - Effectively reduces cost of survey data collection

1. New sources of auxiliary data

- Significant work using geospatial data from satellites
    - Many useful indicators are freely available
    - Geographically comprehensive
    - Not subject to selection bias
    - Updated frequently
    - Very geographically granular,
        - Can often link surveys at EA level or small (admin-4) admin level
        - Can very occasionally link surveys at household level
    - A second-best option to recent census data
        - Census data are richer but often old, don’t reflect recent shocks
    - Extensive and innovative literature (Newhouse, 2024)

- Other non-traditional data intriguing
    - For example, social media, call detail records, and crowd-sourced information
    - Quality varies depending on context

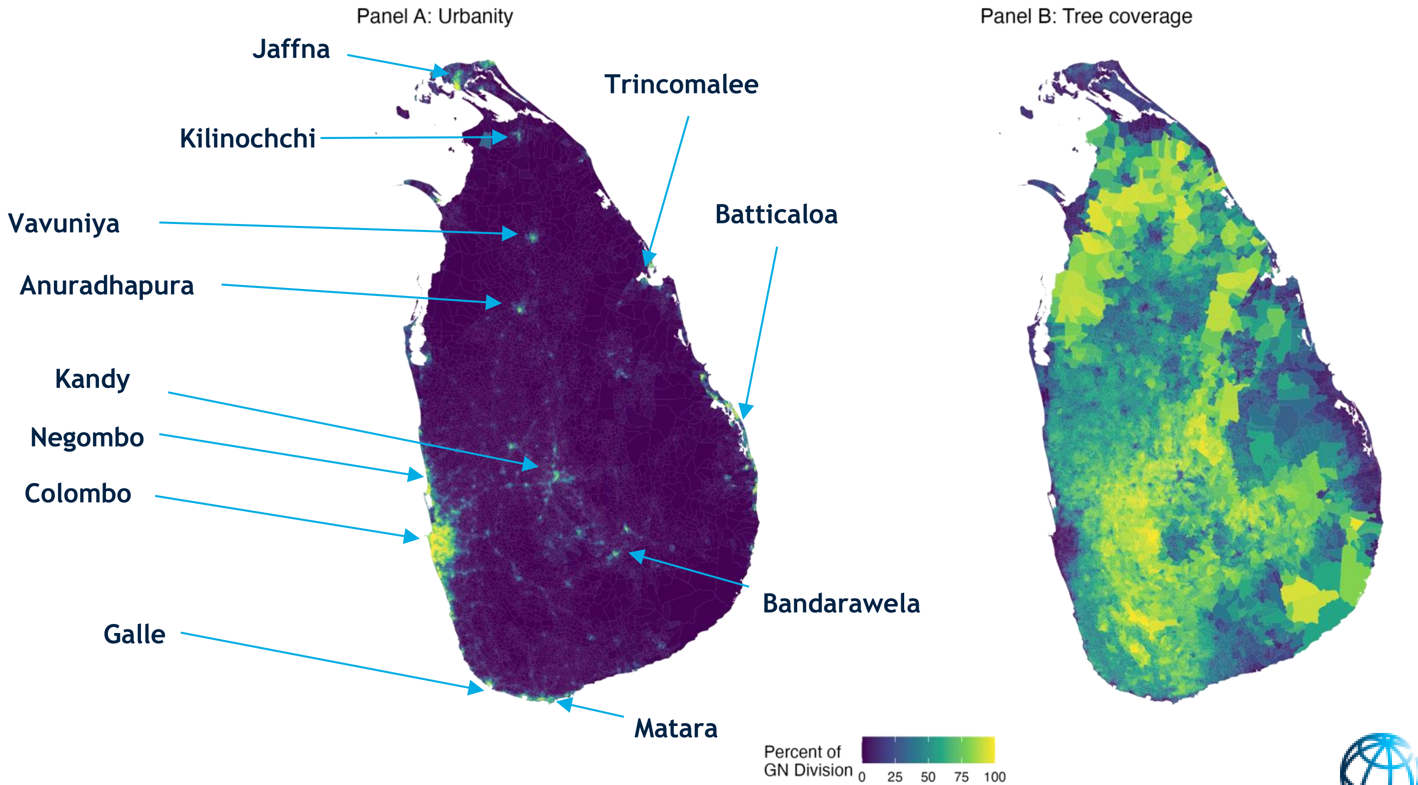
Malawian building footprint data in Microsoft planetary computer



WORLD BANK GROUP

Land classification percentages by GN Division in Sri Lanka

- Clear negative relationship between urbanity and tree coverage
- Urban measure identifies cities and towns

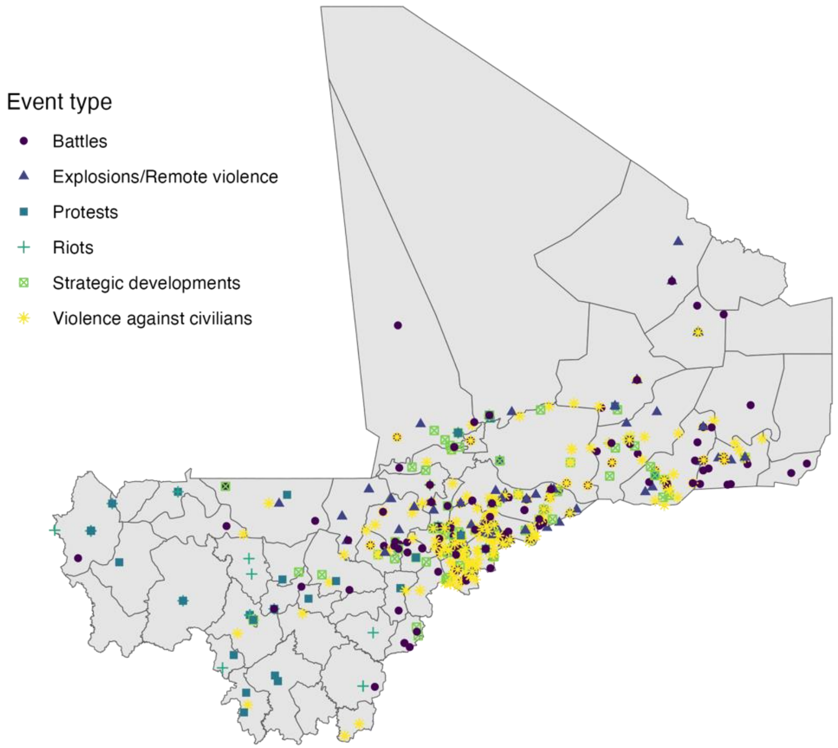


WORLD BANK GROUP

Malian Armed Conflict Location and Events Data (ACLED)

Contains information on location of different types of violent events

Violent events in Mali, 2019



WORLD BANK GROUP

Sample publicly available geospatial indicators

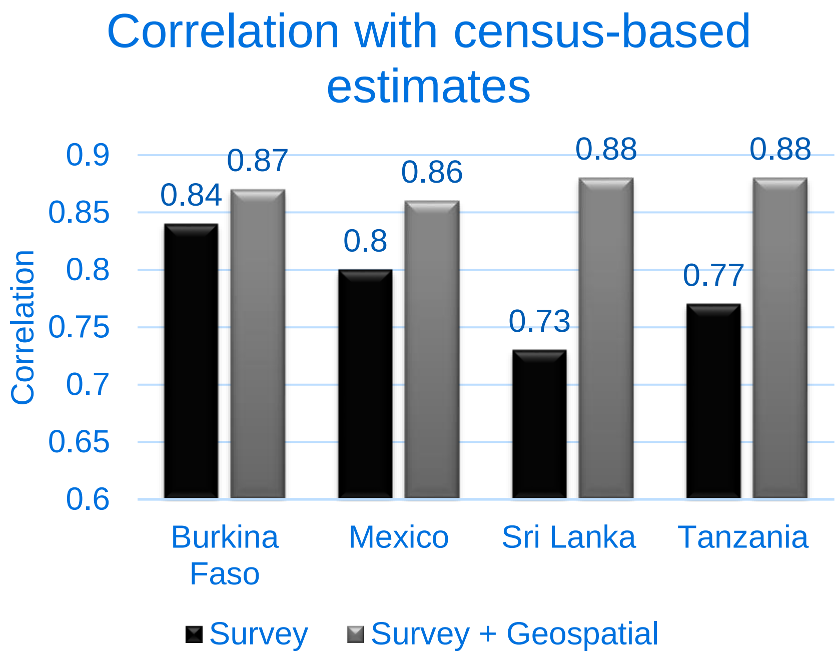
Variable	Source	Resolution	Year
Population structure	WorldPop	100 m	2018
Population density	WorldPop	100 m	2018
Temperature	TerraClimate	4 km	2018
Palmer Draught Severity Index (PSDI)	TerraClimate	4 km	2018
Distance to major roads	WorldPop, Open Streetmap	100 m	2016
Radiance of night-time lights	VIIRS	500 m	2018
Net primary production	FAO Remote Sensing for Water Productivity (WaPOR) 2.1	240 m	2018
Rainfall	Climate Hazards Group InfraRed Precipitation with Station data (CHIRPS)	5.5 km	2018
Elevation	NASA's SRTM Digital Elevation (3 arc seconds spatial resolution)	30 m	2018
Cellphone tower count	The OpenCell ID project	1 km	2022
Years since change to impervious surface	Tsinghua University via Google Earth Engine	30 m	2018
Building count	Worldpop	100 m	2018
Coefficient of variation on buildings	Worldpop	100 m	2018
Land cover classifications	Dynamic World	100 m	2018

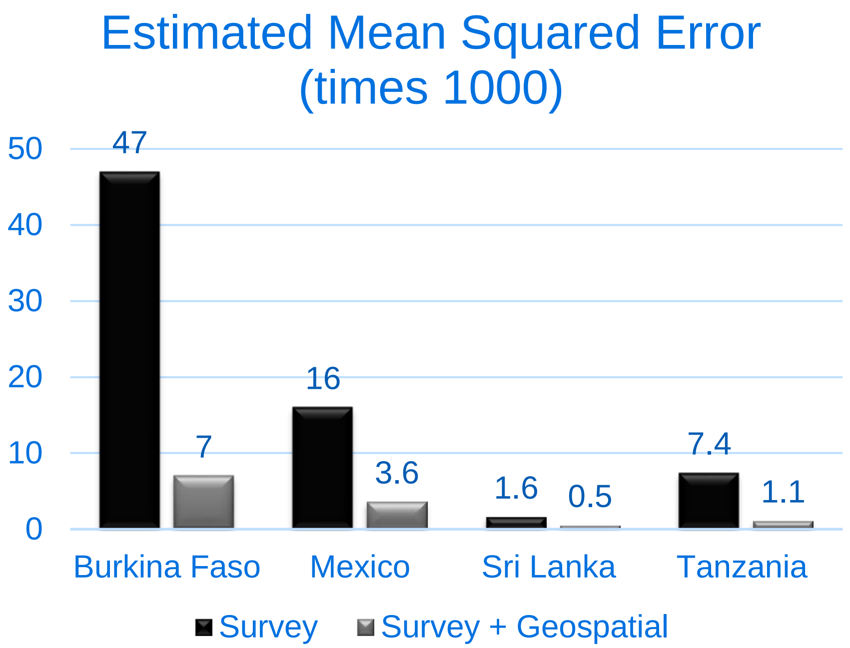
Geospatial indicators are very predictive of population density

- In Sri Lanka, out of sample R^2 of 0.75 when predicting population density with publicly available indicators
  - Increases to 0.83 when using proprietary indicators (Engstrom et al, 2020)
- In DRC, out of sample R^2 0.79 for predictions of population totals at the microcensus-cluster level
  - Boo et al, 2022
- Population density is correlated to many important socioeconomic outcomes
  - Including poverty and wealth (Page and Pande, 2018, Casteneda et al, 2018)
- SAE reduces sampling error at the expense of introducing model error
  - Does this improve accuracy relative to direct estimates?
  - Depends on outcome, sample, and predictive power of auxiliary data

Geospatial small area estimates of poverty are more accurate than survey estimates when tested

- Reflects strong relationship between poverty and population density
- Enables large increases in precision while maintaining coverage





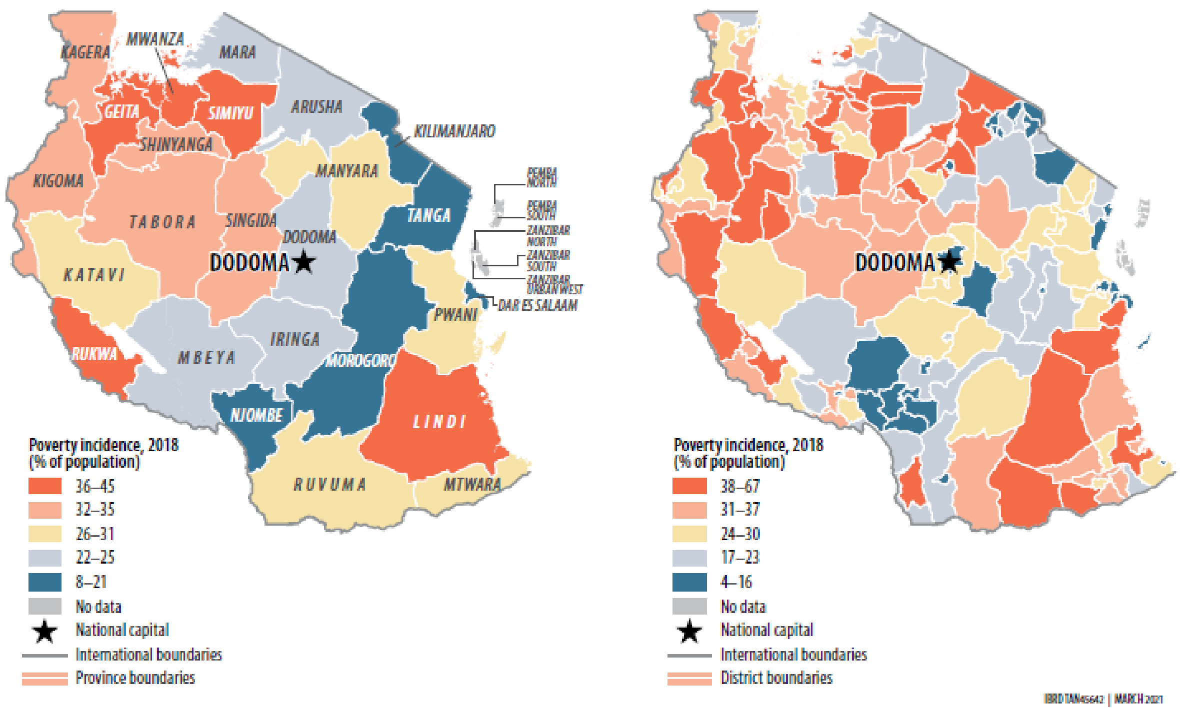
Sources: Masaki et al (2022), Newhouse et al (2023), Edochie et al (Forthcoming)
Notes: Results based on actual household survey data. Survey estimates are direct estimates, survey + geospatial are EBP estimates using a linear mixed model.

Can sometimes enable more disaggregated reporting with same level of precision

Map O.3 Combining satellite imagery with household survey data increases the resolution of the poverty map of Tanzania

a. Poverty map using the Household Budget Survey (20 regions)

b. Poverty map combining the data in panel a with satellite imagery (169 districts)

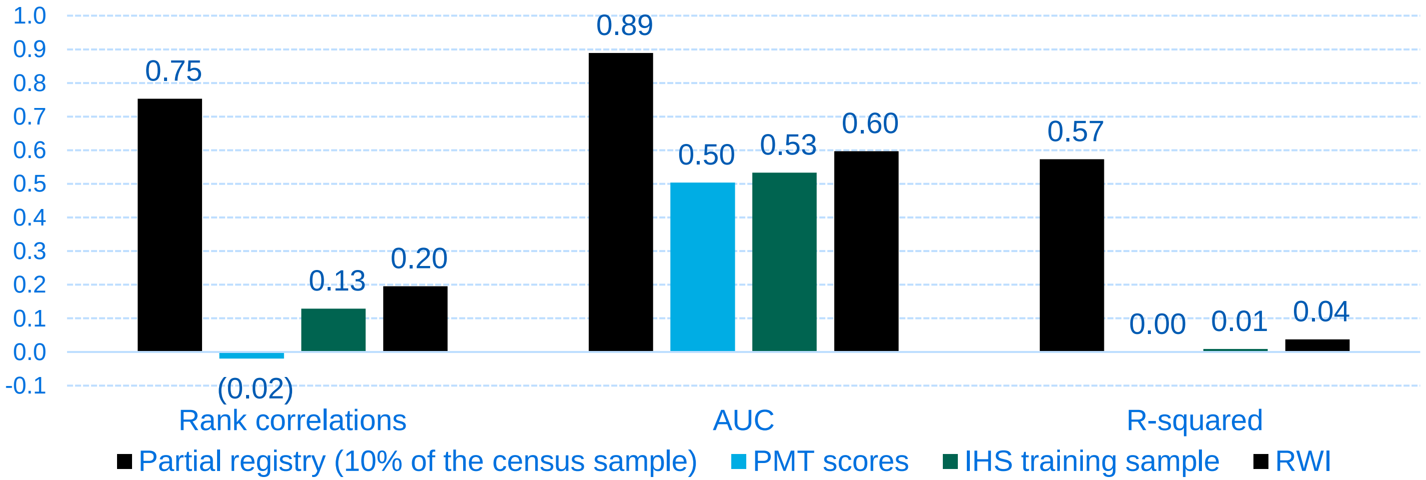


Source: World Bank 2019. Data at http://bit.do/WDR2021-Map-O_3.

WORLD BANK GROUP

Accurate micro-estimates require rich training data

- Predicting village level welfare measures is difficult
  - Not enough data in a typical household survey to train a good model
- Adding data from a partial registry fills the gap and greatly improves accuracy
  - Consider a hypothetical partial registry that collects 15 proxy welfare questions from all households across 450 villages
  - When we simulated this prediction performance improved tremendously
  - Partial registries or other sources of rich training data can unlock the potential of geospatial data for very granular prediction



Source: Gualavisi and Newhouse (2022)

2. Improved statistical methods for SAE

- SAE Methods can be applied to all types of auxiliary data

- Several statistical methods have been proposed
  - Brilliant statisticians and computer scientists constantly proposing improvements
  - Even large language models like ChatGPT (Manvi et al, 2023)

- Need to test, evaluate, verify, and validate new methods in realistic settings
  - Design-based simulations with census data critical for this (Tzavidis et al, 2018)
  - Understand strengths and weaknesses of different models in different contexts
  - Nuanced understanding to guide method selection and help navigate tradeoffs along different dimensions (Das and Haslett, 2019)
  - Methods differ not only in accuracy but also simplicity, interpretability, and robustness (Efron, 2020)

- Incorporate desirable options into user-friendly software

Examples of methodological improvements

1. Improve estimation of EBP models with sample weights
   - Estimating EBP models with weights is complicated
   - There is room for improvement in both commonly used methods

2. Allow for more heterogeneity in EBP models.
   - Can be important when there are a large number of heterogeneous target areas that are systematically different
   - New methods better capture heterogeneity in the relationship between outcome and predictors across areas (Lahiri and Salvati, 2024)

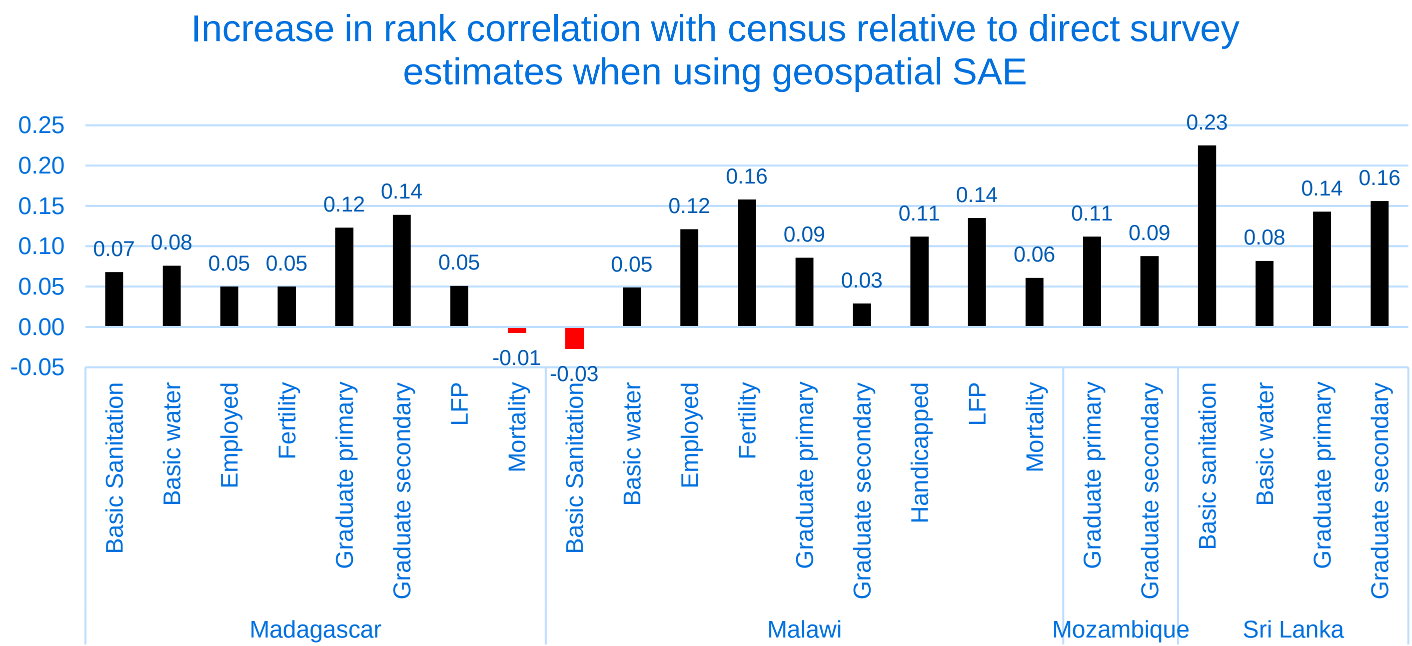
3. Use methods that are more robust to selection bias in sample
   - Different methods are more or less robust to biased samples (Cho et al, 2024, Pfeffermann and Sverchkov, 2007)

4. Better diagnostics
   - How to tell without census data how accurate estimates will be?

3. New indicators beyond poverty

New auxiliary data opens new doors

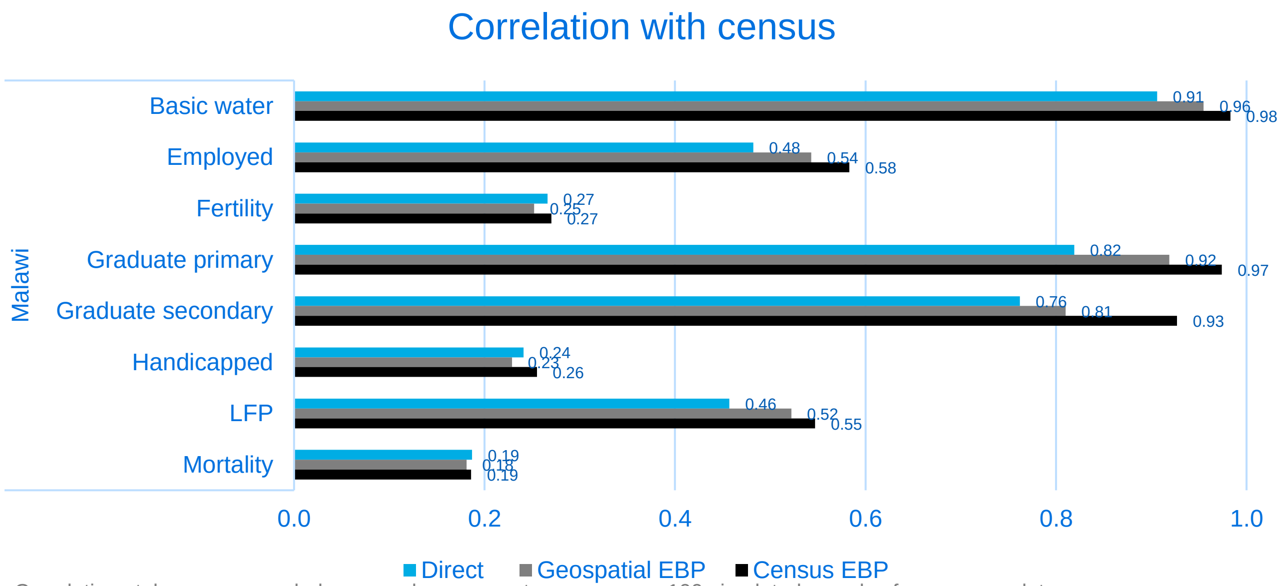
- Geospatial small area estimation can improve accuracy and precision of many labor, health, education, population outcomes from surveys
- Amount of improvement in accuracy varies greatly across outcomes and contexts



WORLD BANK GROUP

Some outcomes are better suited for SAE than others

- In some cases, geospatial data are a partial and useful substitute for census data
  - Basic water, employed, educational attainment, and labor force participation
- For other indicators, SAE doesn't help even with a census
  - Fertility, handicapped, mortality not estimated accurately even with census
- Good models need accurate sample data
- Need to develop better diagnostics to assess accuracy without a census



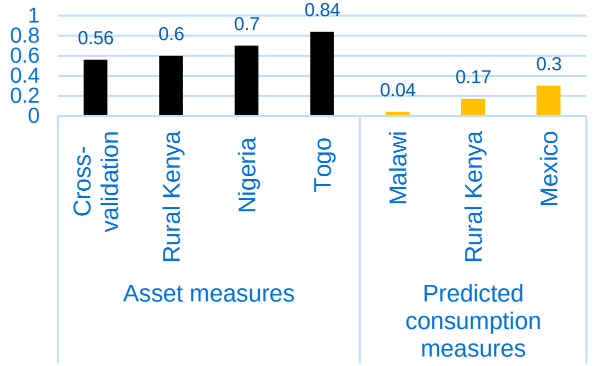
Note: Correlations taken over sampled areas only, represent average over 100 simulated samples from census data

WORLD BANK GROUP

The outcome measure also matters

- For example, Meta relative wealth index is much more accurate when measuring asset poverty than consumption poverty
  - Wealth used as outcome because of data availability
  - Consumption or income is standard measure for poverty measurement

R^2 from evaluation against benchmarks



Sources: Chi et al (2022), Gualavisi and Newhouse (2023), Newhouse et al (2022)

- Illustrates importance of implementing SAE in National Statistics Offices
  - Providing access to geospatial data and tools to NSOs and others with sensitive survey data
  - NSOs can model outcomes while preserving data privacy
  - Well-documented, user-friendly, and freely accessible tools are key

4. Tools and documentation are essential

1. Geolink R package
   - A “one-stop shop” that will download geospatial data, process it, and merge it with survey data
   - Links either though latitude/longitude coordinates or through administrative identifiers with shapefiles
   - Indicators will include night-time lights, land classification, modeled population estimates, Open Streetmap points of interest, rainfall, temperature
   - Aiming for release in late 2024

4. Povmap

2. Povmap for small area estimation
    - Extends EMDI (Estimating Mapping and Disaggregated Indicators) function which was built on R SAE package
    - Version 1.0 released last September
    - Version 2.0 in production

- Advantages of EMDI over alternatives
    - Very well-documented and user-friendly
    - Implements sample and population weights
    - Estimates both unit-level and area-level models with recent innovations
        - Wide variety of transformations including adaptive transformations for improved accuracy
    - Automates and facilitates basic diagnostics
    - Parallelization of MSE estimation for increased speed
    - Fully open-source

4. Povmap

Povmap extends EMDI by adding options for:

- Benchmarking to ensure consistency with survey estimates at higher levels
  - Properly accounting for uncertainty
- Alternative method for incorporating sample weights
  - Preferable in some cases
  - Adds options for automatic weight rescaling
- Additional transformations
- Helpful pre and post-estimation commands for summarizing results
- More control over details of model estimation
- Stata wrapper command to for Stata users

4. Povmap 2.0 in development

- Intend to support additional methods
  - Twofold model (Marhuenda et al, 2018)
  - “ELL method” (Elbers, Lanjouw, and Lanjouw, 2003)
  - Extreme gradient boosting (Chen and Guestrin, 2016)
  - High-dimensional parameters (Lahiri and Salvati, 2024)

- Support for expected value calculation for means and headcounts
  - Much faster than repeated simulations and slightly more accurate

- Intend to offer additional weighting options

- Intended release: Late 2024

Conclusions

- Rapid recent advances in publicly geospatial auxiliary data
    - Clearly useful in some important cases
    - Population, poverty, wealth, labor force participation, ag?, Others?
    - Second-best alternative to recent available census data
    - Not well-suited for all outcomes (but neither is census data)
    - Stock of publicly available geospatial indicators should continue to improve
    - Granular prediction can benefit greatly from additional training data beyond standard surveys

- Need to continue pushing on methodological modernization
    - Test, Evaluate, Validate, and Verify new proposed methods
    - Better understand strengths and weaknesses of different methods in different contexts
    - Improve diagnostics

- New applications
    - Beyond poverty mapping
    - Data augmentation can improve accuracy and precision of survey statistics at all geographic levels

- Continue to develop tools and documentation to democratize access and build capacity

Conclusions

- Consultation draft of “Small Area Estimation with Geospatial Data: A Primer” (joint with UNSD) available at https://unstats.un.org/iswghs/

Table of Contents

WORLD BANK GROUP

References

Barrington-Leigh, C., & Millard-Ball, A. (2017). The world's user-generated road map is more than 80% complete. PloS one, 12(8), e0180698.

Biljecki, F., Chow, Y. S., & Lee, K. (2023). Quality of crowdsourced geospatial building information: A global assessment of OpenStreetMap attributes. Building and Environment, 237, 110295.

Battese, G. E., Harter, R. M., & Fuller, W. A. (1988). An error-components model for prediction of county crop areas using survey and satellite data. Journal of the American Statistical Association, 83(401), 28-36.

Boo, G., Darin, E., Leasure, D. R., Dooley, C. A., Chamberlain, H. R., Lázár, A. N., ... & Tatem, A. J. (2022). High-resolution population estimation using household survey data and building footprints. Nature communications, 13(1), 1330.

Carter, G. M., & Rolph, J. E. (1974). Empirical Bayes methods applied to estimating fire alarm probabilities. Journal of the American Statistical Association, 69(348), 880-885.

Castañeda, A., Doan, D., Newhouse, D., Nguyen, M. C., Uematsu, H., & Azevedo, J. P. (2018). A new profile of the global poor. World Development, 101, 250-267.

Chen, T., & Guestrin, C. (2016, August). Xgboost: A scalable tree boosting system. In Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining (pp. 785-794).

Chi, G., Fang, H., Chatterjee, S., & Blumenstock, J. E. (2022). Microestimates of wealth for all low-and middle-income countries. Proceedings of the National Academy of Sciences, 119(3), e2113658119.

Cho, Y., Guadarrama-Sanz, M., Molina, I., Eideh, A., & Berg, E. (2024). Optimal Predictors of General Small Area Parameters Under an Informative Sample Design Using Parametric Sample Distribution Models. Journal of Survey Statistics and Methodology, smae007.

Corral, P., Molina, I., Cojocaru, A., & Segovia, S. (2022). Guidelines to small area estimation for poverty mapping.

Das, S., & Haslett, S. (2019). A comparison of methods for poverty estimation in developing countries. International Statistical Review, 87(2), 368-392.

Fay III, R. E., & Herriot, R. A. (1979). Estimates of income for small places: an application of James-Stein procedures to census data. Journal of the American Statistical Association, 74(366a), 269-277.

Edochie, I., D. Newhouse, T. Schmid, and N. Wurz, "Povmap: Extension to the emdi package for small area estimation", available at Comprehensive R Archive Network

Edochie, I., D. Newhouse, T. Schmid, N. Tzavidis, E. Foster, A. Ouedraogo, A. Sanoh, and A. Savadogo (forthcoming), "Small area Estimates of Poverty in Four West African countries", mimeo

Efron, B. (2020). Prediction, estimation, and attribution. International Statistical Review, 88, S28-S59.

Efron, B., & Morris, C. (1977). Stein's paradox in statistics. Scientific American, 236(5), 119-127.

References

Elbers, C., Lanjouw, J. O., & Lanjouw, P. (2003). Micro-level estimation of poverty and inequality. Econometrica, 71(1), 355-364.

Engstrom, R., Newhouse, D., & Soundararajan, V. (2020). Estimating small-area population density in Sri Lanka using surveys and Geo-spatial data. PloS one, 15(8), e0237063.

Gualavisi, M., & Newhouse, D. L. (2022). Integrating Survey and Geospatial Data to Identify the Poor and Vulnerable, Policy Research Working Paper no. 10257.

Jiang, J., & Lahiri, P. (2006). Mixed model prediction and small area estimation. Test, 15, 1-96.

Lahiri, P., & Salvati, N. (2023). A nested error regression model with high-dimensional parameter for small area estimation. Journal of the Royal Statistical Society Series B: Statistical Methodology, 85(2), 212-239.

Marhuenda, Y., Molina, I., Morales, D., & Rao, J. N. K. (2017). Poverty mapping in small areas under a twofold nested error regression model. Journal of the Royal Statistical Society Series A: Statistics in Society, 180(4), 1111-1136.

Manvi, R., Khanna, S., Mai, G., Burke, M., Lobell, D., & Ermon, S. (2023). Geollm: Extracting geospatial knowledge from large language models. arXiv preprint arXiv:2310.06213.

Masaki, T., Newhouse, D., Silwal, A. R., Bedada, A., & Engstrom, R. (2022). Small area estimation of non-monetary poverty with geospatial data. Statistical Journal of the IAOS, 38(3), 1035-1051.

Molina, I., & Rao, J. N. (2010). Small area estimation of poverty indicators. Canadian Journal of statistics, 38(3), 369-385.

Newhouse, D., Merfeld, J., Ramakrishnan, A. P., Swartz, T., & Lahiri, P. (2022). Small Area Estimation of Monetary Poverty in Mexico using Satellite Imagery and Machine Learning.

Newhouse, D. (2023). Small Area Estimation of Poverty and Wealth Using Geospatial Data: What have We Learned So Far?. Calcutta Statistical Association Bulletin, 00080683231198591.

Page, L., & Pande, R. (2018). Ending global poverty: Why money isn't enough. Journal of Economic Perspectives, 32(4), 173-200.

Pinheiro, J., & Bates, D. (2006). Mixed-effects models in S and S-PLUS. Springer science & business media.

Tzavidis, N., Zhang, L. C., Luna, A., Schmid, T., & Rojas-Perilla, N. (2018). From start to finish: a framework for the production of small area official statistics. Journal of the Royal Statistical Society Series A: Statistics in Society, 181(4), 927-979.

WORLD BANK GROUP

References