Use of Alternative Sources and Techniques for the Integration of Statistical and Geospatial Information.

Segment 2:

“Geospatial information for Sustainable Development”

August 18, 2021

El futuro es de todos

Gobierno de Colombia

UN-GGIM
UNITED NATIONS
COMMITTEE OF EXPERTS ON
GLOBAL GEOSPATIAL
INFORMATION MANAGEMENT

DANE
INFORMACIÓN PARA TODOS

Contents

UN-GGIM
UNITED NATIONS
COMMITTEE OF EXPERTS ON
GLOBAL GEOSPATIAL
INFORMATION MANAGEMENT

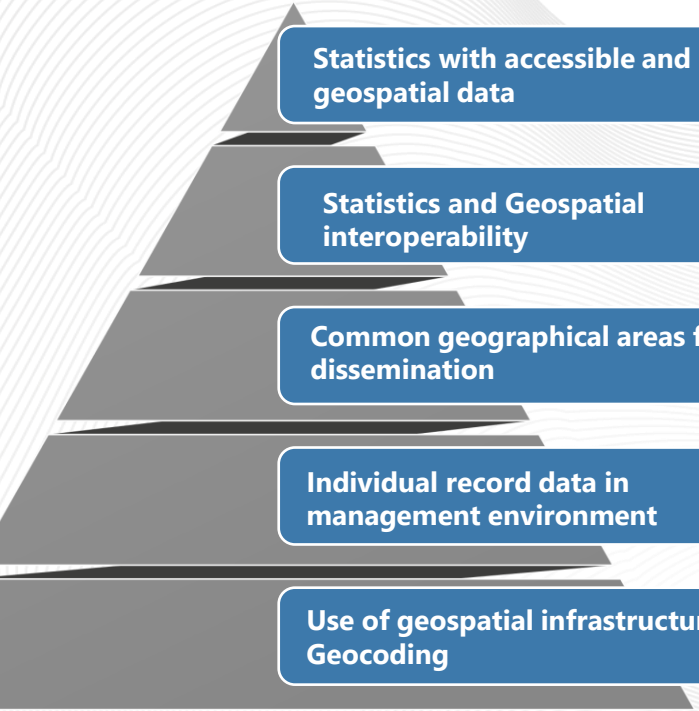
DANE
INFORMACIÓN PARA TODOS

1. DANE's foundations for the integration of statistical and geospatial information

El futuro es de todos
Gobierno de Colombia

Global Statistical and Geospatial Framework

Five general principles in the integration of statistical and geospatial information



Statistics with accessible and usable geospatial data	**Discovery of Statistics** by promoting services for **visualization and analysis.**
Statistics and Geospatial interoperability	Increase data standardization and interpretation for efficiency and better simplification.
Common geographical areas for statistics dissemination	Establish definitions of geographic regions and the aggregation/break down of data.
Individual record data in management environment	Connect each individual or statistical record to a **geographic reference.**
Use of geospatial infrastructure and Geocoding	**Develop infrastructure that allows the statistical and geospatial interoperability.**

https://ggim.un.org/meetings/GGIM-committee/9th-Session/documents/The_GSGF.pdf

Key Aspects

Currently, the integration of statistical and geospatial information is essential for :

1. Local, sub-national , national, regional and global **decision making processes**
2. **Measuring and monitoring SDG targets and indicators**
3. Supporting **data sharing** between institutions
4. **Creating new knowledge** and linkages between diverse sources of information
5. Promoting investment and capacity development
6. **Examining new data sources**

Integration of statistical and geospatial information for the calculation of SDG indicators

Three key benefits of including Earth observations in the NSO's task of calculating SDG indicators

I. The possibility of **deriving SDG indicators**, which would be technically and financially difficult to calculate.

II. Decrease the frequency of surveys and the associated costs for **providing information in high level of disaggregation.**

III. **Provide breakdown and granularity of indicators**, ensuring that they are spatially oriented.

2. Use of alternative sources and techniques for the integration of geospatial information for statistical production

Production of experimental statistics

Calculation of SDG indicator 1.2.2 *Multidimensional Poverty Index (MPI)*
Census MPI prediction using machine-learning and satellite imagery

MPI Dimensions

- Housing standards and access to public utilities
- Childhood and youth conditions
- Health
- Employment opportunities and conditions
- Access to education

In line with → SUSTAINABLE DEVELOPMENT GOALS

✓ **Need** for periodic poverty measurements

Background
- Bangladesh
- Ghana
- Uganda

Inputs

- Census MPI
- Census Indicators
- MGN

- Person
- Home
- Housing

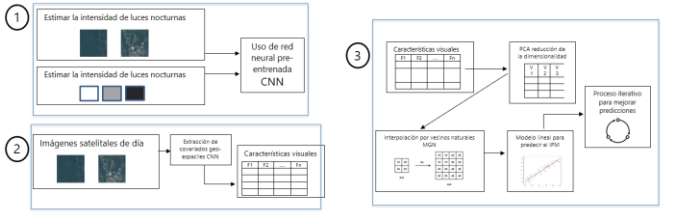
- Sentinel-2 images (77,979 images)
- Nighttime light intensity

https://www.dane.gov.co/index.php/estadisticas-por-tema/estadisticas-experimentales

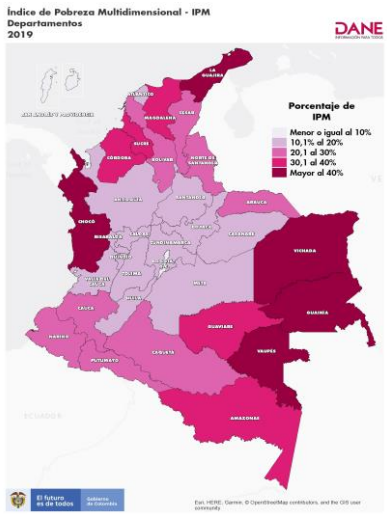
INFORMACIÓN PARA TODOS

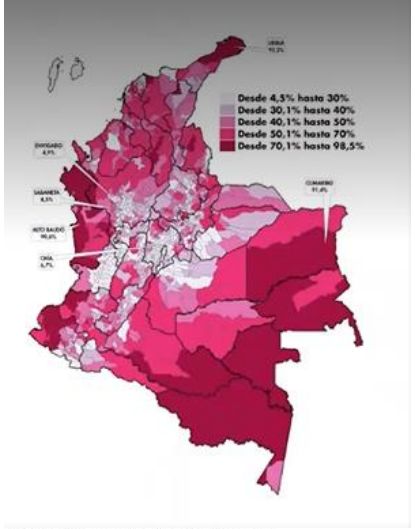
INFORMACIÓN PARA TODOS

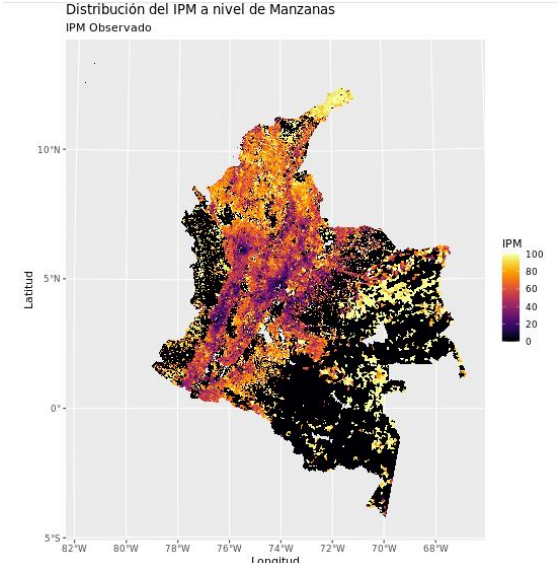
Proceso de transferencia de aprendizaje para predecir el IPM a partir de imágenes satelitales

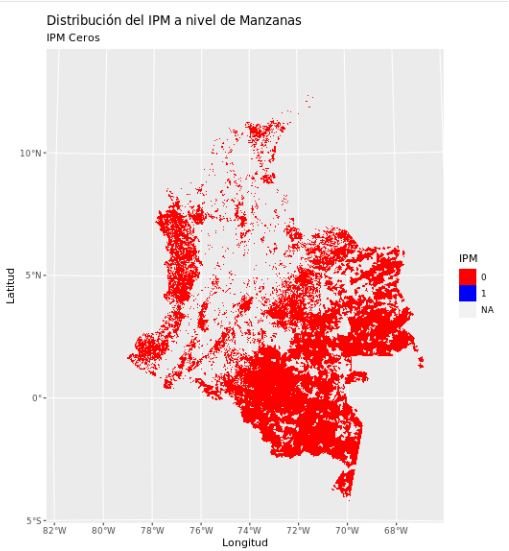


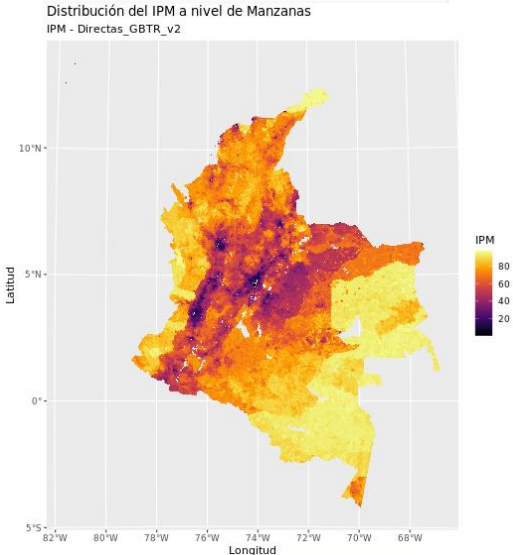
1) Se usa una *Convolutional Neural Network (CNN)* para predecir la intensidad de la luz nocturna.
2) Las características visuales de alto nivel se extraen de las capas superiores de la CNN.
3) Se entrena un modelo lineal y se mejora a partir del uso de un proceso iterativo de *Gradient Boosting Tree Regression models*











Production of experimental statistics

Calculation of SDG indicator 9.1.1

Proportion of rural population living within 2 km of a road that is passable all year round.

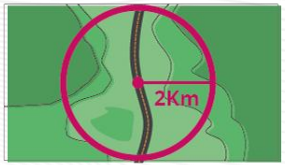
Universe of study:
Calculation of the indicator for all departments with final information.

Sources:
- 2018 National Housing and Population Census .
- Official cartography of the Agustín Codazzi Geographic Institute

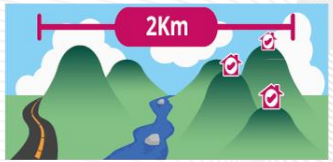
Summary of the Methodology:

Step 1
Determination of year-round roads (source IGAC)

Step 2
Methodological development based on the World Bank's Rural Accessibility Index, plus the inclusion of impedances (hydrography - elevations).



Step 3
Calculation of the area of influence of 2 km from the previous result.



Step 4
Intersection with the geo-referencing of dwellings, to determine the rural population within the area of influence.

Step 5
Calculation of the indicator from the previous results

https://www.dane.gov.co/index.php/estadisticas-por-tema/estadisticas-experimentales

Production of experimental statistics

Calculation of SDG indicator 11.3.1

Ratio of land consumption rate to population growth rate.

11 SUSTAINABLE CITIES AND COMMUNITIES

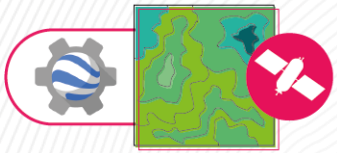
Universe of study:

Cities defined through the degree of urbanization methodology (DEGURBA).

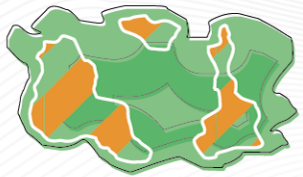
Sources:

- Satellite imagery for determination of land consumption.
- Population projections to calculate the growth rate.

Summary of the Methodology:



Step 1
Selection and classification of satellite images



Step 2
Identification of the built-up area in the 63 defined cities

Step 3
Determination of the area of change in land consumption

Step 4
Determination of the relationship between land consumption rate and population projections 2015 - 2020.

Step 5
Calculation of the indicator from the previous results

https://www.dane.gov.co/index.php/estadisticas-por-tema/estadisticas-experimentales

Production of experimental statistics

Calculation of SDG indicator 11.7.1

Average share of the built-up area of cities that is open space for public use for all, by sex, age and persons with disabilities

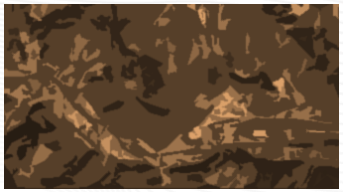
Universe of study:
National sample of cities based on the methodology suggested by the UN-Habitat Global Urban Observatory

Sources:
- Satellite images to determine the built area.
- Open access sources - citizen generated data: Open Street Map (OSM)
- Georeferenced statistical information (CNPV 2018)

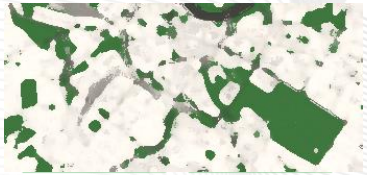
Summary of the Methodology:

Step 1
Definition of a national sample of cities with the methodology of the Global Urban Observatory (GUO)

Step 2
Debugging of sources and integration of inputs



Step 3
Defining built-up areas through Sentinel-2 image classification



Step 4
Identification of open spaces for public use (green areas, roads)

Step 5
Calculation of the indicator in the cities of the sample, national estimate and disaggregation by population groups

Production of experimental statistics

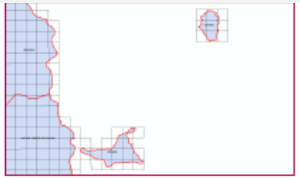
Ethnic population estimation at sub-national levels

Phases:

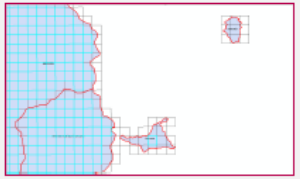
1. Determine the participation of people self-recognized as belonging to an ethnic group at the national level.
2. Assign ethnic affiliation to individuals.
3. Determine the participation of each ethnic group at levels of disaggregation below the municipal level, based on **the georeferencing of the census population**.
4. Distribute proportionally the omitted population not assigned in phase 2 and 3, by ethnicity, through the Square Table methodology.

✓ **Estimated the 2018 NHPC omission by indigenous reservation fragment.**

Circumscribed Grid

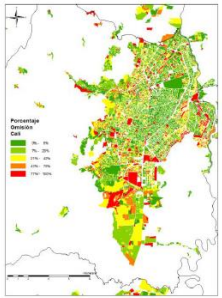


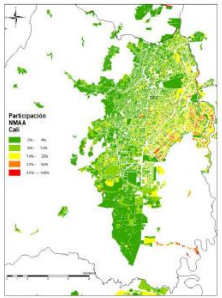
Enrolled Grid

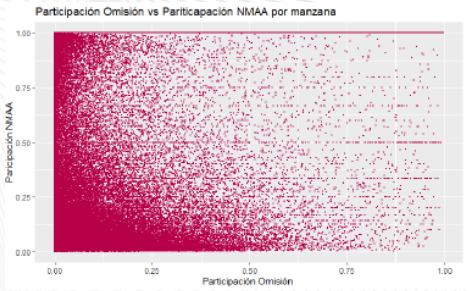


Grid Fragment

✓ **Determination of percentage of omission and census participation of the NMAA population at the per block or rural sector level.**







https://www.dane.gov.co/index.php/estadisticas-por-tema/estadisticas-experimentales

Updating and upkeep of statistical frameworks

Rural and Agricultural Master Framework

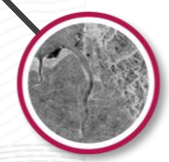
Methodological proposals for the continuous updating of area frameworks through the use of satellite images and drones.

Strengthening the Agricultural Framework

First Line

- Identification of **domain crops** exploring new classification methods, based on the project developed in 2020 (potatoes and beans).

Formulate a **methodology for the identification and estimation of specific crop areas** (potatoes and beans).



Second line

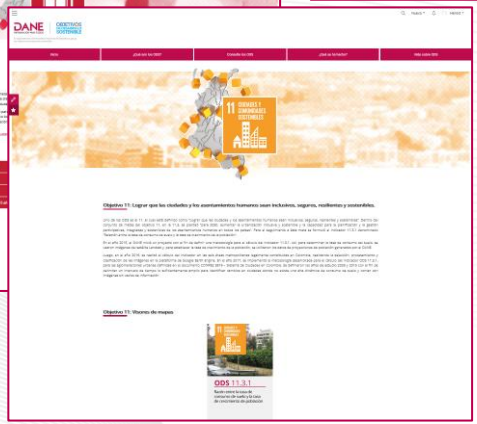
- Identification of coverages in **difficult to access areas**, with different inputs (radar) and classification methods.

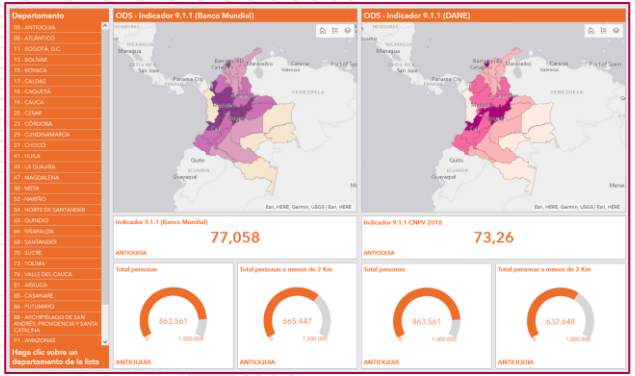
Formulate a **methodology for land cover classification in areas that are difficult to access.**

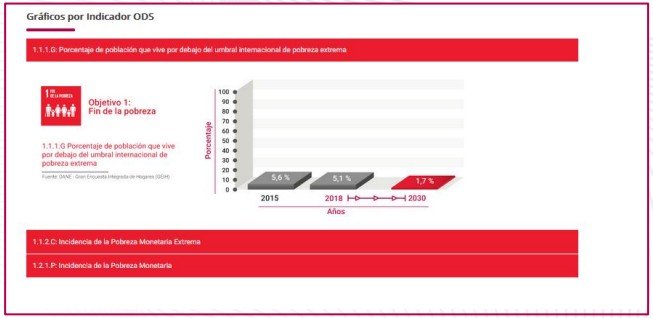
Result Dissemination

ArcGIS Hub for the dissemination of SDG indicators

Portal developed for the presentation of results, through graphics and geo-viewers in infrastructure provided by Esri







https://indexods-ec702-dane-ods.opendata.arcgis.com/

Use of Alternative Sources and Techniques for the Integration of Statistical and Geospatial Information.

Segment 2:

“Geospatial information for Sustainable Development”

August 18, 2021

El futuro es de todos

Gobierno de Colombia