FIG Working Week 2024
19-24 May
Accra, Ghana
Your World, Our World: Resilient Environment and Sustainable Resource Management for All



# Use of Sentinel 2 Images and Cadastral Data to Analyze Urban Growth in Colombia

Sonia Constanza Garzón, Rodolfo Yesid Ramírez Patacon
and Johan Andrés Avendaño Arias

Paper 12677

Dirección de Investigación y Prospectiva/Directorate of Research and Perspective
Instituto Geográfico Agustín Codazzi (IGAC)

Mayo 2024/May 2024

IGAC
INSTITUTO GEOGRÁFICO AGUSTÍN CODAZZI

# WHAT DO WE DO?

**We produce the official map and cartography of Colombia**

**We train and educate professionals in geographic information technologies**

**We are the Cadastral authority**

**We carry out geographical investigations**

**We carry out an inventory of the characteristics of the soils**

**We coordinate the Colombian Spatial Data Infrastructure (ICDE)**

## Research and Prospective technical direction

- Research Center recognized. SNCTeI .
- CIAF. More than 50 years of experience in research, development and training in geospatial technologies. 25.000 students from Colombia and the Latin American region.

**R&D**

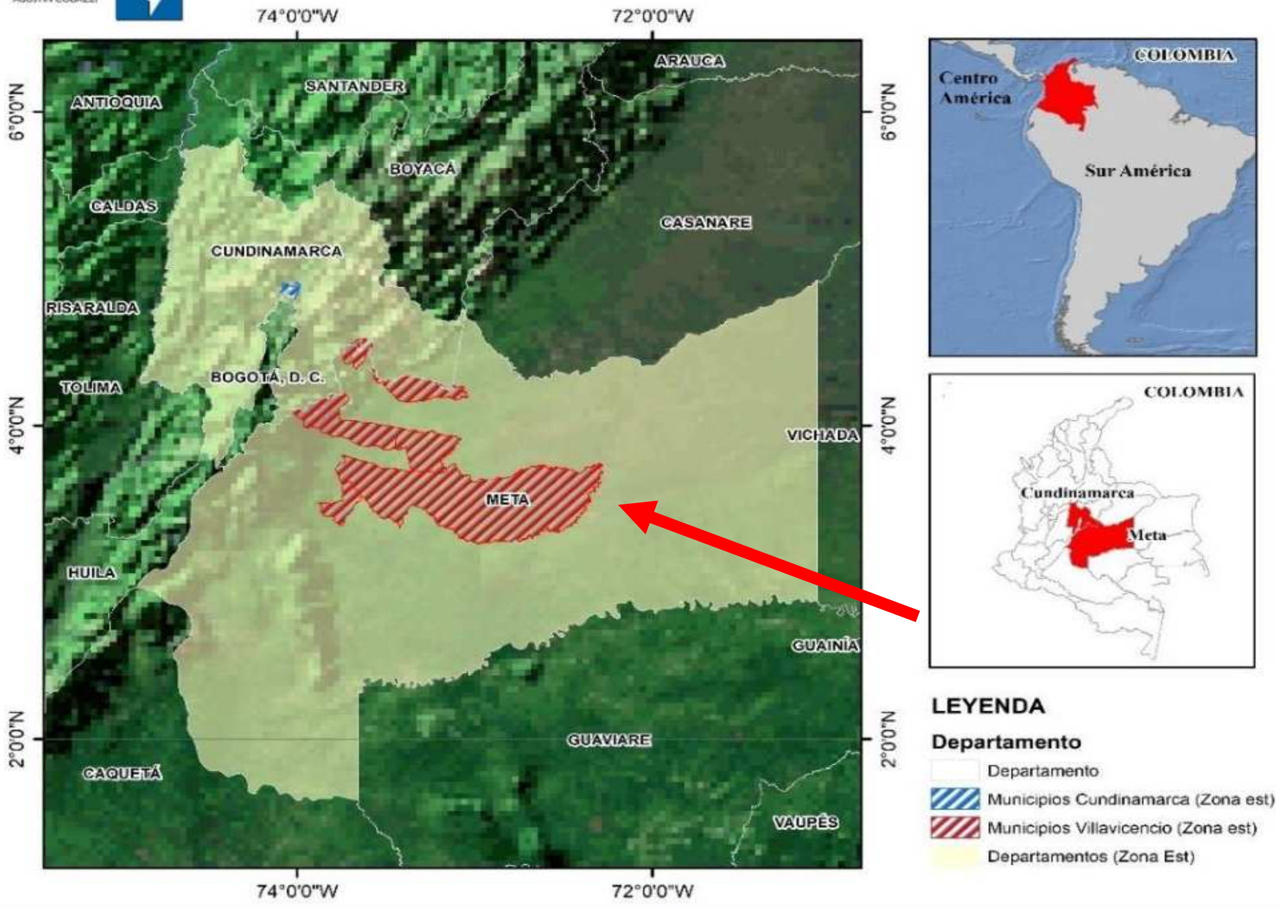
**Prospective**

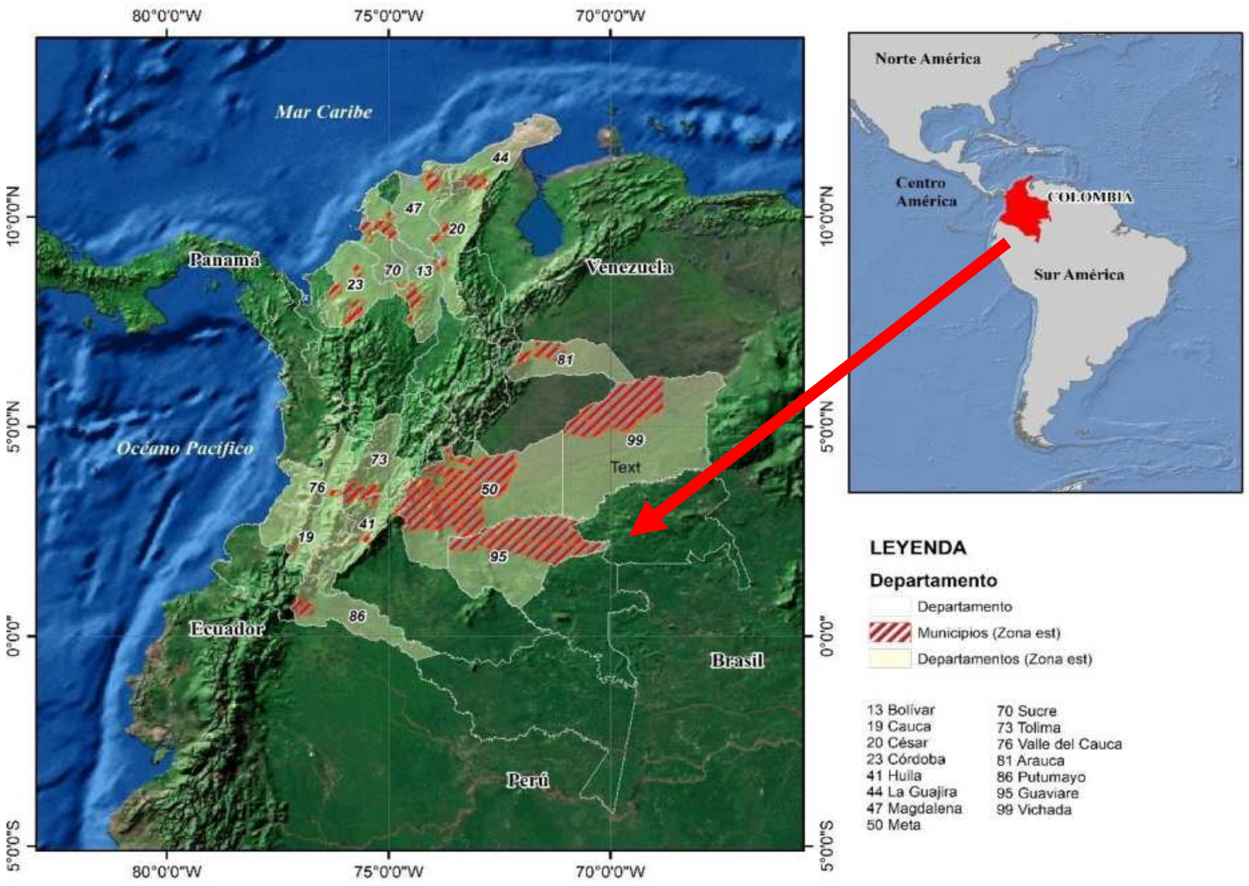
**Cadastral Real Estate Dynamics**

**Training and qualification**

IGAC INSTITUTO GEOGRÁFICO AGUSTÍN CODAZZI

# Area of study:



IGAC
INSTITUTO GEOGRÁFICO AGUSTÍN CODAZZI

# Introduction:

Since 2016, Colombia has experienced an elevated level of outdated cadaster, with only 320 out of 1101 municipalities having cadastral updates. Pilot exercises in 11 municipalities revealed limited field work inputs, causing delays in cadastral surveys.

The IGAC was assigned to review technologies for identifying urban expansion, property changes, and construction changes as indicators of territory dynamics in cadastral matters. In 2020 and 2022, the IGAC advanced AI models for classification of high-resolution images and comparing construction levels between different periods.

A convolutional neural network model was also reviewed to predict urban expansion. By 2022, the IGAC faces challenges in achieving a 70% update rate from 9.4% in 2022 to 70% by 2026, as per the National Development Plan (2022-2026).

**Key Words:** Urban growth, Sentinel 2, Artificial Intelligence, Data Analysis, and Cloud Computing.

IGAC
INSTITUTO GEOGRÁFICO AGUSTÍN CODAZZI

## Objectives:

Evaluate, adapt and scale the results of the models generated to different regions of the country, especially in those areas prioritized by the cadastral updating process.

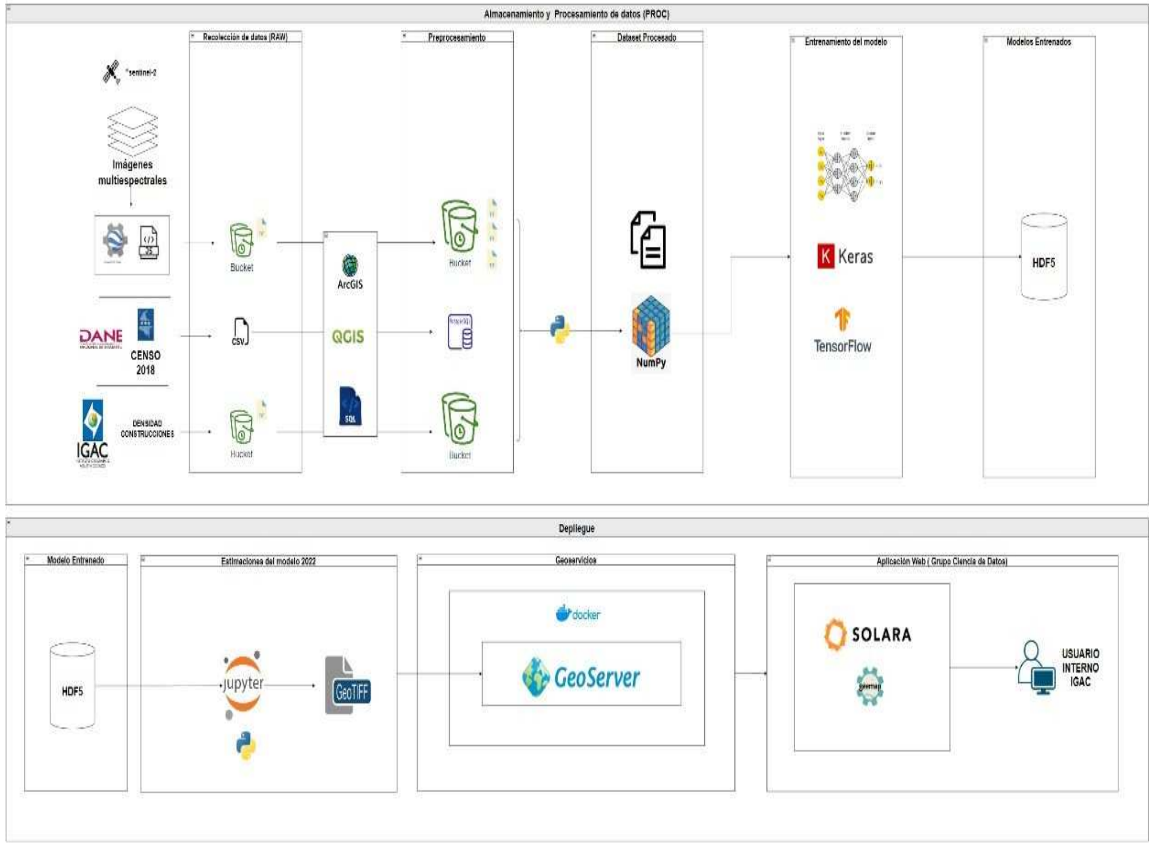
Take advantage of institutional and open-source data that allow implementing the new technological solution as part of the productive process of the cadastral management advanced by the IGAC and define a process of extraction, transformation and ETL loading linked to the different data sets.

Define the architecture of a system for monitoring physical changes for cadastral purposes, capable of modularly coupling the different components that make up the indicators of change in the territory.

# Methodology (Inputs):

IGAC
INSTITUTO GEOGRÁFICO AGUSTÍN CODAZZI

- **Sentinel-2 imagery:** The basis of the analysis was the Sentinel-2 satellite imagery acquired through the Google Earth Engine API. These images are fundamental for the identification and monitoring of land cover changes and urban sprawl.
- **Sentinel-2 Image Preprocessing:** The areas corresponding to each block were extracted from Sentinel-2 satellite imagery, with special attention to the 2018 census data. This step was crucial to ensure that the model was trained with geographically accurate and up-to date information.
- **DANE 2018 Census Data:** Another key component in the model was the 2018 Census data provided by the National Administrative Department of Statistics (DANE). The use of these inputs, combined with advanced data processing techniques, allowed the creation of a robust and reliable model for the analysis of urban sprawl.



IGAC INSTITUTO GEOGRÁFICO AGUSTÍN CODAZZI

# Methodology (Tools):

- The urban expansion study in the department of Meta, Colombia, relied on a diverse set of technological tools for data processing and model training. These tools facilitated large-scale data analysis and allowed for greater accuracy and efficiency throughout the process.
- The use of Jupyter Notebooks provided an interactive and user-friendly platform for experimentation and data visualization. The integration of ipywidgets into the notebooks further enhanced the interactivity and user experience, allowing for more dynamic and accessible data analysis.

- Google Earth Engine (GEE).
- Python and Jupyter Notebooks.
- Integration and Scalability.

Google Earth Engine

Python

jupyter

IGAC
INSTITUTO GEOGRÁFICO AGUSTÍN CODAZZI

## Methodology (Step by Step)

### Information Processing

- Data loading
- Image processing (GEE)
- Dataset construction

### Model training

- Train and Test split
- Train and optimization

### Solution deployment

- Solution scalability
- Data download optimization
- Integration od Tools and PLatforms
- Solution Architecture Definition

IGAC
INSTITUTO GEOGRÁFICO AGUSTÍN CODAZZI

# Results (Physical change monitoring system architecture)

- The quality and relevance of the data used in the training plays a key role in the results, in this case the labels are calculated from information available from the DANE 2018 Census selected for its inherent relationship to urban growth (Aspinall, R, 2004).
- Considering the presence of missing data in each of these characteristics, a composite indicator is constructed that involves proportionally each of the selected census characteristics and reduces the possibility of assigning labels with null values in the training.
- Final data sets were constructed with Numpy array format, ensuring the elimination of null data and the proper standardization of values between 0 and 1 to optimize the learning process and the convergence of the neural network.
- These arrays are constructed from tessellations associated with 13x13px regions, with center at the pixel labeled with the ICU, following Tobler's first law (Sui, D, 2004).

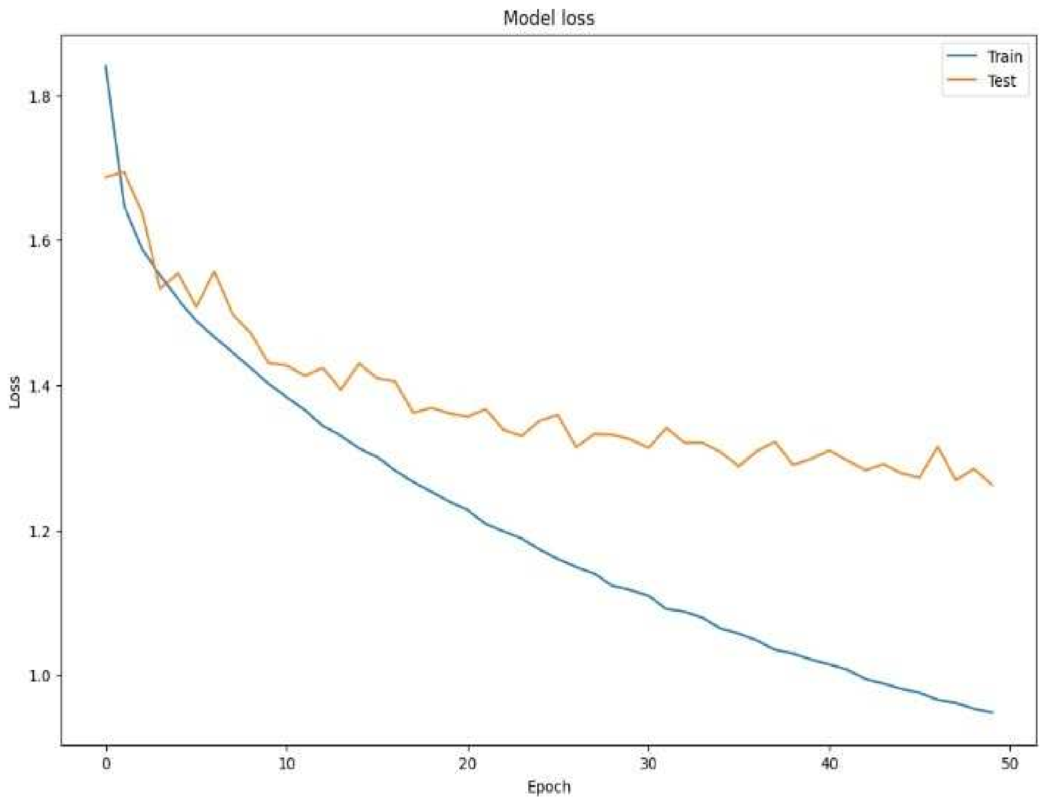
$$ICU = \sum_{J=1}^{pxmanzana} k1 * POBLACION + K2 * VIVIENDA + K3 * DENSIDAD$$

**Composite urban growth index.**

IGAC INSTITUTO GEOGRÁFICO AGUSTÍN CODAZZI

# Results (Evaluation, adaptation, and scaling of the model)

- Model training was conducted in Chia, Colombia to detect urban growth using convolutional neural networks and satellite images.
- Despite the limited study area, the initial phase proved the feasibility of the implementation. Manual steps were taken in constructing the dataset with the aid of GIS software and Python scripts.
- The goal of the subsequent phase was to scale the model by expanding the data processing and training to different municipalities.
- Computational capacity was enhanced and data augmentation was employed to improve model generalization.
- The results from a 10GB dataset encompassing 6 municipalities in Meta demonstrated satisfactory learning behavior with both the training and test data. However, the test data showed higher loss.



# Results (Physical change monitoring system architecture)

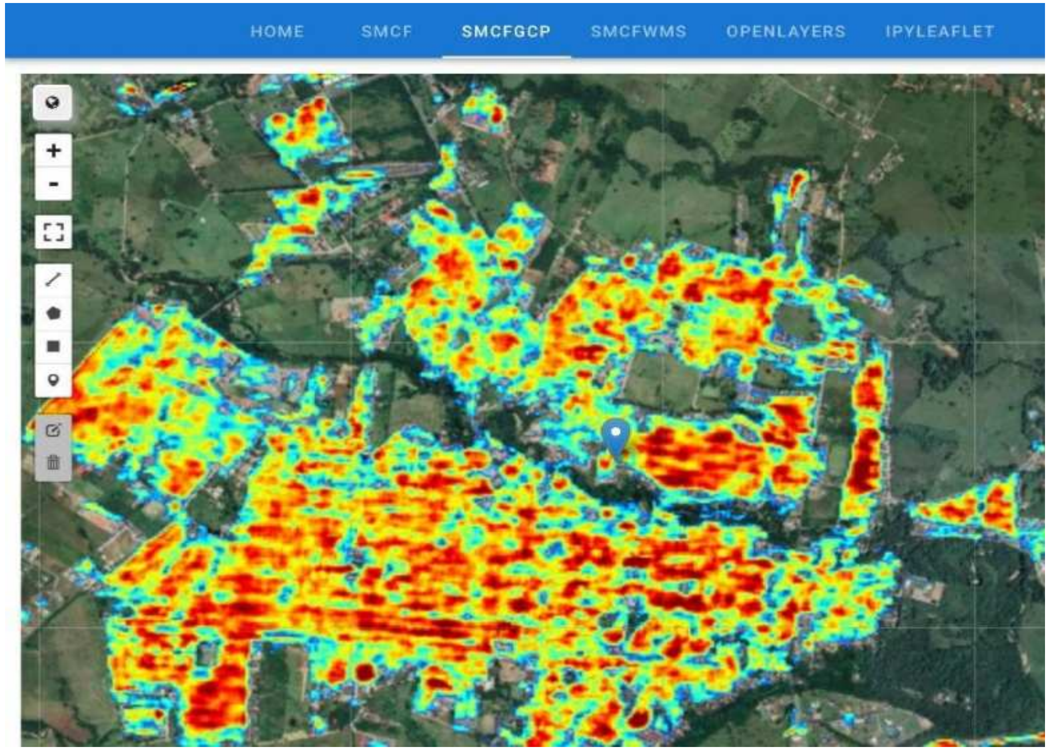
IGAC INSTITUTO GEOGRÁFICO AGUSTÍN CODAZZI

- The monitoring system we have developed to study physical changes in urban areas in Colombia is structured in two main blocks: information processing and model deployment. This architecture is designed to efficiently capture, process, and visualize data related to urban sprawl.

1). The first stage of the system focuses on geospatial data processing and model building. This includes the collection and analysis of satellite imagery, census data, and other relevant sources.

2). The second phase of the system involves deploying the model and visualizing the results.

3). Finally, for interactive visualization of the results, we employed the Solara framework and Geemaps in a web application.



IGAC INSTITUTO GEOGRÁFICO AGUSTÍN CODAZZI

## Conclusions:

- The Sentinel project's satellite images, despite having lower spatial resolution compared to Google Earth or Bing, are adequate for detecting significant urban growth variations.
- Some changes in the images may be due to factors other than urban growth, such as canopy changes or weather conditions.
- The model was initially trained with data from 6 municipalities in Meta and later with 49 prioritized municipalities, but the results showed too much variation to be satisfactory.
- Other approaches: Future research, Model Training Improvements, Working with more Complex Data Sets, Integration of New Data Sources and Technical and Scalability Challenges.

SUSTAINABLE DEVELOPMENT GOALS

International Federation of Surveyors supports the Sustainable Development Goals

# Commission #

## Commission’s name

Serving Society for the Benefit of People and Planet

# Thank you all !!

**LADM_COL (ESP)**

**LADM_COL (ENG)**

**Johan Andrés Avendaño Arias**
Technical Director of Research and Prospective
Geographic Institute Agustín Codazzi
johan.avendano@igac.gov.co

**Authors:**
**Johan Andrés Avendaño Arias:** johan.avendano@igac.gov.co
**Sonia Constanza Garzón :** sonia.garzon@igac.gov.co
Rodolfo Yesid Meza: rodolfo.meza@igac.gov.co

**Intercultural School of Geography for Life**