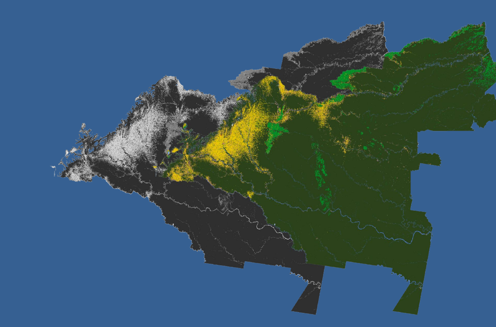
Geospatial Data Analytics for Policy Impact Assessment: Amazon Deforestation in Colombia

Earlham College

Michelle Marchesini Vanegas - mmarch22@earlham.edu

Computer Science Capstone '26



Introduction

The Amazon Rainforest is a protected area that spans across 9 countries, covering around 1% of the planet's surface. In Colombia, the armed conflict has exacerbated forest loss as rival groups compete for land control in coca-growing regions, pushing cultivation deeper into the forest. The 2016 Peace Agreement has aimed to alleviate this issue, but questions remain about whether post-conflict measures have effectively reduced deforestation rates.

This capstone project addresses this critical gap by assessing the effectiveness of post-conflict conservation policies and legal frameworks. By combining satellite remote sensing with temporal policy analysis, this research provides a data-driven evaluation of how well current interventions are translating into measurable forest outcomes.

GOALS

- **Develop automated geospatial data processing pipelines** to efficiently ingest, classify, and analyze multitemporal satellite imagery using machine learning algorithms, enabling scalable monitoring of forest cover.
- **Integrate heterogeneous datasets** through geospatial data fusion techniques, combining satellite-derived land cover classifications with modeled scenarios and socio-political event timelines to create a unified analytical framework for policy-outcome correlation analysis.
- **Build temporal analysis systems** that automatically identify inflection points in deforestation trends and correlate them with policy interventions enabling evidence-based assessment of conservation mechanism effectiveness.
- **Create reproducible, open-source computational workflows** that can be adapted and scaled to monitor deforestation across the broader Colombian Amazon and transnational Amazonian regions, supporting future conservation monitoring and decision-making.

MATERIALS AND METHODS

This study integrates two primary satellite-derived datasets to assess deforestation across temporal and spatial scales:

Dataset 1: Historical Land Cover (2001–2016)
- Landsat-derived annual maps via NASA EarthData
- Generated using CCDC algorithm & Random Forest classifier
- 8 land cover classes identified

Dataset 2: Future Projections (2003–2050)
- SimAmazonia model scenarios
- Business-as-Usual & Governance scenarios
- 50% deforestation cap & Protected Area enforcement

Land Cover Classification

A supervised Random Forest classifier trained on manually collected reference data distinguished eight land cover categories:

Import GeoTIFF files imported into QGIS for spatial processing	**Visualization** Color-coding applied to distinguish land cover classes and identify spatial patterns
Quantification Pixel counting analysis to determine percentage coverage and area extent	**Analysis** Temporal and socio-political integration for policy assessment

Integrated Analysis Approach

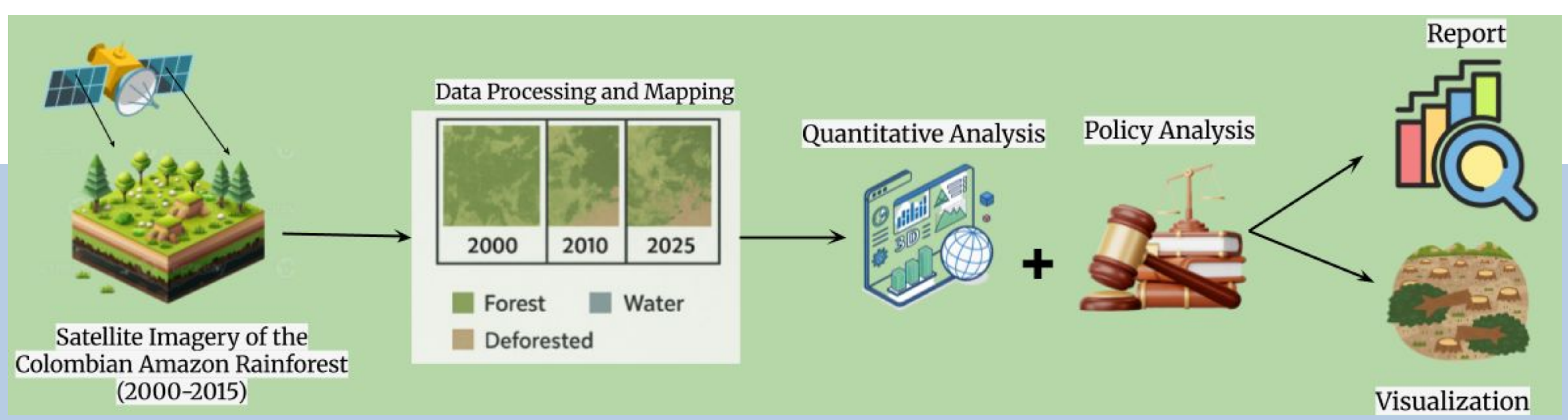
Temporal Mapping: Deforestation metrics plotted chronologically alongside socio-political events (armed conflicts, policy changes, protected area establishment)

Inflection Point Detection: Identification of periods where deforestation accelerated, decelerated, or stabilized relative to policy interventions

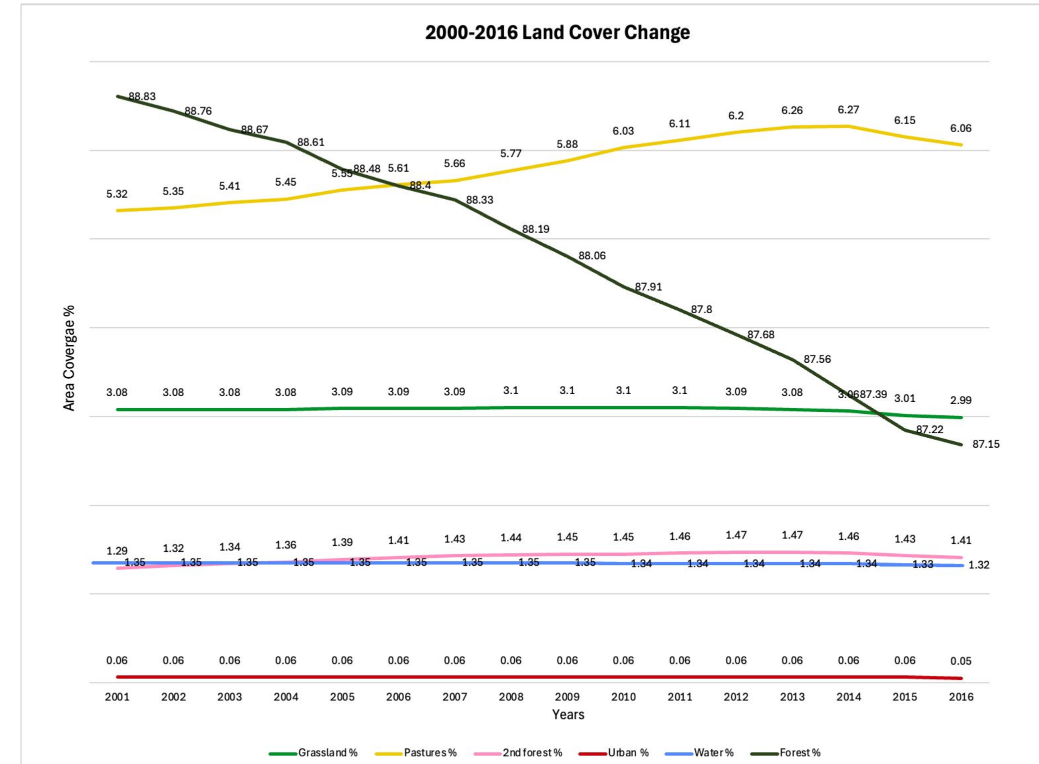
Policy-Outcome Correlation: Assessment of whether specific conservation measures or post-conflict transitions corresponded with measurable changes in forest protection

Regional Variation: Recognition that national-level policies manifest differently across geographical areas with varying state presence and conflict history

Core Innovation: This methodology bridges remote sensing and policy analysis by integrating satellite-derived deforestation metrics with temporal socio-political context. Rather than treating deforestation as a purely technical phenomenon, this approach enables evidence-based assessment of how legal frameworks, peace processes, and conservation mechanisms translate into measurable forest outcomes



RESULTS



From 2000 to 2016, forest cover shows a steady decline while pastures expand, indicating a clear conversion of forested land into agricultural use over time. Other land-cover classes remain relatively stable, suggesting that deforestation is primarily driven by pasture expansion rather than urban growth or changes in water or grassland areas.

Business-as-Usual Scenario	Governance Scenario
• **Forest loss:** 58,981 km² • **Deforestation increase:** +57,374 km² • Continuous acceleration of forest clearing • Reflects absence of effective intervention • Assumes historical trends continue with road expansion	• **Forest loss:** 41,231 km² (30% less than BAU) • **Deforestation increase:** +41,231 km² • Deforestation rates level off in later years • Assumes 50% deforestation cap per subregion • Protected Areas effectively limit forest conversion

- **Key Finding:** Policy-driven governance mechanisms can effectively constrain deforestation and preserve approximately 17,750 km² of additional forest area over the 47-year projection period (2003–2050), demonstrating the protective potential of conservation-oriented policies and protected area enforcement.
- **Model Validation:** While the NASA SimAmazonia model accurately captures deforestation trends, it systematically underestimates observed forest loss between 2003–2023, particularly after the mid-2010s. This suggests the model's limited sensitivity to episodic drivers such as policy shifts, market pressures, and environmental events. Recent observed deforestation rates exceed model projections, indicating the need for recalibration with contemporary data

DISCUSSION

This study contributes to Amazonian deforestation research by incorporating the socio-political conditions that shape land-use change, while also demonstrating the role of computer science in environmental analysis. Beyond traditional ecological approaches, this work applies computational modeling, geospatial data processing, and scenario-based simulations to analyze deforestation under different governance conditions. While much of the existing literature focuses on Brazil, the results highlight how armed conflict, territorial control, and post-conflict governance dynamics influence forest loss in Colombia. Integrating computer science methods with socio-political context enables more scalable, reproducible, and policy-relevant analyses, supporting evidence-based conservation strategies that account for both technological insights and human systems

ACKNOWLEDGMENTS

I gratefully acknowledge NASA EarthData and Global Forest Watch for providing the data resources that made this analysis possible. I thank the SimAmazonia modeling team for generating the satellite-derived land cover maps and deforestation scenarios used in this study. I also appreciate the support of Earlham College and my faculty advisors for their guidance throughout this geospatial analysis project. This work contributes to broader efforts to evaluate conservation policies and peace-building initiatives aimed at protecting Colombia's Amazon rainforest.

FUTURE DIRECTIONS

- **Integrate causal variables** (road networks, land tenure, military zones) to better attribute deforestation to specific drivers.
- **Develop near real-time monitoring systems** that automatically flag emerging deforestation hotspots for rapid intervention.
- **Create web dashboards** for policymakers and communities to visualize deforestation trends and support evidence-based decision-making.
- **Implement deep learning models** (CNNs) to improve land cover classification and anomaly detection in challenging conditions.

SOURCES

[15] Paula Areevalo. CMS: Landsat-derived Annual Land Cover Maps for the Colombian Amazon, 2001–2016,2020. Accessed December 13,2025.

[16] Britaldo S. Soares-Filho, Daniel C. Nepstad, Lisa M. Curran, Elmar Voll, Ricardo A. Garcia, Claudia A. Ramos, Andrew J. McDonald, Paul A. Lefebvre, and Peter Schlesinger. LBA-ECO LC-14 Modeled Deforestation Scenarios, Amazon Basin: 2002–2050, 2013. Accessed December 13, 2025.