

Dataset List: 2024 Grantee Convening

Lacuna Fund: Our Voice in Data

Climate & Energy, Climate & Health, and Climate & Forest Domains

SEPTEMBER 2024
NEW YORK CITY, USA

DATASET LIST

* = individual is participating in 2024 Lacuna Fund Grantee Convening

CLIMATE & ENERGY DATASETS

20-Month Dataset on Household Electricity Consumption and its Drivers, Collected via Meter Readings, and Longitudinal Survey

Climate & Energy (2022)

Team Members: Merl Syrex Chandana (LIRNEasia),* Nilusha Kapugama

Primary Institution: LIRNEasia

Partner Institutions: Lanka Electricity Company

Project Description: The dataset integrates 20 months of electricity consumption data with results from a three-round longitudinal survey, designed to capture the drivers, attitudes, and knowledge related to electricity use. This data will cover over 4,000 households, selected from the customer base of Sri Lanka's second-largest electricity distribution licensee, which serves approximately 500,000 households along the Western coastal belt of the country.

Addressing AI Challenges: This project demonstrates the power of AI to address climate challenges by leveraging data to optimize energy consumption. Through supervised and unsupervised machine learning techniques applied to the dataset, interested parties could analyze household electricity usage patterns to identify inefficient behaviors and target interventions. These interventions, which might include personalized recommendations, appliance buy-back programs, or tailored pricing plans, can help reduce overall energy demand and emissions, especially in a country like Sri Lanka heavily reliant on fossil fuels for power generation.

Artificial Intelligence for Energy Access

Climate & Energy (2022)

Team Members: Catherina Cader (Reiner Lemoine Institut)*

Primary Institution: Reiner Lemoine Institut (RLI)

Partner Institutions: Abdou Moumouni University – WASCAL, Atlas AI, Clean Technology Hub

Project Description: The AI4EA project, lead by the Reiner Lemoine Institute (Germany), focuses on pioneering solutions to enhance electricity access in West Africa. Partnering with institutions like Atlas AI PBC (USA), Université Abdou Moumouni (Niger), and the Clean Technology Hub (Nigeria), this initiative leverages artificial intelligence to develop high-resolution electricity demand profiles for non-urban regions of Niger, Nigeria, Benin, Togo, and Ghana.

Addressing AI Challenges: The project utilizes comprehensive datasets gathered across multiple geopolitical zones in Nigeria through PeopleSuN project at RLI, alongside advanced remote sensing data from Atlas AI and additional regional datasets provided by WASCAL. This data informs the development of robust AI models that are carefully trained, validated, and tested to ensure their accuracy and reliability. Based on this data, clean renewable energy-based supply solutions can be developed sustainably.

Dataset for Energy consumption, needs & forecasting using Indigenous tools

Climate & Energy (2022)

Team Members: Gul Muhammad Khan (CISNR)*, Zeeshan Shafiq, Atif Zeeshan

Primary Institution: Centre for Intelligent Systems and Networks Research (CISNR) UET Peshawar

Partner Institutions: Sarhad Rural Support Program

Project Description: The dataset comprises of Generation data in the form of Voltages, currents, power factors, energy generated, active and reactive power (all electricity parameters) of microhydal power plants, distribution transformers and consumers. Thus covering a complete eco system starting from generation to ultimate consumption at the end user. The data of two power houses, comprises of four generators, providing electricity to over 4000 consumer using 56 transformers, although through this grant we managed to implement the system on 1000 users, and 30 transformers only.

Addressing AI Challenges: AI is used to forecast the demand and generations, and balancing it by load management to ensure quality energy access to the end users. The data can also be used to predict the highly fluctuating water channels and demand statistics due to climate change causing an increase in frequency of floods ultimately effecting both demands and generation.

Development of Emission Dataset for Abattoir Operation facilities in Southern Parts of Nigeria

Climate & Energy (2022)

Team Members: Emmanuel Chukwuma (Alliance for Progressive and Sustainable Environment)*

Primary Institution: Alliance for Progressive and Sustainable Environment

Partner Institutions: McGill University, Nnamdi Azikiwe University

Project Description: The dataset is data on air quality from air quality sensors and drone images from aerial views of the study area.

Addressing AI Challenges: The major cause of climate change is anthropogenic activities, through the mitigation of the impact of abattoir operations in the study area, carbon emission can be reduced

Distributed Energy Resource Management Dataset

Climate & Energy (2022)

Team Members: Sunimal Rathnayake (University of Moratuwa),* Chirath Pathiravasam

Primary Institution: University of Moratuwa

Partner Institutions: Lanka Electricity Company, Magicbits Ltd

Project Description: A dataset consisting of renewable energy generation data from from different sources such as solar inverters as well as corresponding energy demand data from appliances such as air conditioners.

Addressing AI Challenges: This dataset enables minimizing energy cost, minimizing system peak, maximizing reliability and maximizing consumer comfort via AI based predictions that lead to sustainable energy consumption.

EDADES: Ecuadorian Dataset on Access, Demand, and Availability of Electricity Supply Climate & Energy (2022)

Team Members: Jose Cordova-Garcia (ESPOL University)*

Primary Institution: ESPOL University

Partner Institutions: CENACE

Project Description: EDADES stands for Ecuadorian Dataset on Access, Demand, and Availability of Electricity Supply. This project aims to consolidate data from different sources to make it electrical energy information more accessible. EDADES also includes data collected directly from users to create a proxy for access and availability, particularly important in the context of climate change, that is often overlooked.

Addressing AI Challenges: EDADES will facilitate the development of AI-based applications for detecting failures that affect energy accessibility and that currently are not observable by system operators until users are widely affected. We are working on bringing down time the time to bring these data-driven applications into production from months to hourly resolution updates.

Hyperlocal Mapping of Air Pollution and GHG Emissions in India Climate & Energy (2022)

Team Members: Shubham Tandon (UNDP),* Soumen Maity

Primary Institution: United Nations Development Programme (UNDP)

Partner Institutions: Development Alternatives, Bihar State Pollution Control Board, University of Nottingham

Project Description: Upon project completion, our dataset will include extensive air pollution data from two of India's most polluted cities, which are also among the most polluted in the world. This dataset comprises over 10 million spatial and temporal data points on air quality, along with movement data from approximately 100 citizen scientists who contributed to the project within these cities.

Addressing AI Challenges: When air pollution data is generated using IoT devices and citizen scientists, AI models can provide a comprehensive analysis of pollutant concentrations across the city. This helps to clear identification of the most polluted areas and exploration of potential causes. Different use case scenarios can be developed using these AI models, enabling government agencies to take informed and targeted actions to address climate change challenges effectively.

Open Solar Panel Data Madagascar Climate & Energy (2022)

Team Members: Rafidiharirina Fabienne Mananarivoarisoa (Madagascar Initiatives for Digital Innovation)*

Primary Institution: Madagascar Initiatives for Digital Innovation

Project Description: Renewable energy data from solar panel annotation of satellite and drones images allowing accurate solar panel detection. The project aims to create solar panel location data for future ML detection throughout Madagascar. We want to expand our data collection to different cities of

different climate and geolocation to improve labellisation and enhance quantity to be as representative of the country at its maximum. In Madagascar, Machine learning application and even geolocated data is very rare for most sector and for energy transition, we hope this dataset to be used by the Ministry of Energy to assess the adoption of renewable energy and by journalists for their datajournalism project in the energy and/or environmental sector in Madagascar.

Addressing AI Challenges: This data can facilitate renewable energy detection measuring impact of energy on climate.

Residential Energy and Weather Datasets (REWD) of Pakistan

Climate & Energy (2022)

Team Members: Naveed Arshad (Lahore University of Management Sciences)*

Primary Institution: Lahore University of Management Sciences

Partner Institutions: KU Leuven, National University of Singapore

Project Description: In our dataset we plan to collect detailed residential electrical consumption information from buildings across six climatic zones in Pakistan. The dataset will include data from various areas of the building including floors, rooms and high powered appliances and so on.

Addressing AI Challenges: The data is a good representation of typical buildings in Pakistan and can be utilized to develop AI-based interventions in energy conservation, distribution planning, demand side management, energy forecasting, renewable energy provisioning, energy efficiency and so on.

Solar Irradiance Dataset for Mauritius, Rodrigues and Agaléga

Climate & Energy

Team Members: Dr. Yogesh Beeharry (University of Mauritius)*

Primary Institution: University of Mauritius

Partner Institutions: Mauritius Renewable Energy Agency, kiloWattsol

Project Description: This dataset comprises of Global Horizontal Irradiance (GHI) data collected at a hourly interval from 12 sites across Mauritius, Rodrigues, and Agalega islands. The sites include Agalega, Bramsthan, Curepipe, Ferney, Goodlands, Grenade, La Mivoie, Mahebourg, Nicolay, Port Mathurin, Rose Hill, and Souillac. The data were acquired using base stations consisting of TMF100 Data Loggers and MS-802 EKO Pyranometers. Each site's data collection duration varies. The dataset is stored in CSV format and includes timestamp in the format yyyy-mm-dd HH:MM, site name, latitude, longitude, and Global Horizontal Irradiance measurements (W/m^2).

Addressing AI Challenges: With the ongoing data collection and ML models developed, the local utility will have a tool that helps with the planning for electricity generation using solar energy. This in turn will have a direct impact on the grid with regards to the percentage energy generation using solar energy.

Links and Publications: <https://doi.org/10.57760/sciencedb.08095>

Urban roadspace classification, taxonomy and label training dataset

Climate & Energy (2022)

Team Members: Maria Fernanda Ramirez Bernal (Fundación Despacio),* Don Spencer, Camilo Urbano

Primary Institution: Fundación Despacio

Partner Institutions: World Resources Institute

Project Description: We will annotate street features from aerial photos and cross-examine its qualitative and quantitative data to develop an algorithm that recognizes sidewalks, green spaces, BRT, bus, bike and vehicle lanes, among others. By doing this, it will be possible to quantify how much city space is being taken by each mode.

As a result, we will create a tool to measure road space distribution by using aerial photos, like the ones from Google Maps.

Addressing AI Challenges: This will be a decision-making tool to help cities' leaders to estimate how is road space is distributed, to contrast mode-shift and validate active and sustainable mobility investments and policies.

CLIMATE & HEALTH DATASETS

Climate Change, Built Environment and Health in Informal Settlements in Colombia

Climate & Health (2022)

Team Members: Alex Quistberg (Drexel University)*

Primary Institution: Drexel University

Partner Institutions: Universidad de los Andes, Universidad del Norte

Project Description: The ESCALA (Study of Urban Health and Climate Change in Informal Settlements in Latin America) dataset links longitudinal health risk and outcomes, climate, built and social environment characteristics, and annotated informal settlement characteristics in satellite & street-level imagery in two Colombian cities (Bogota and Barranquilla) at the neighborhood level. Our data have also been informed by community engagement in both cities to understand the perspectives of people living in informal settlements on climate risks and their needs to address those risks, and we are preparing to share the linked data with these communities and local stakeholders to empower them and facilitate conversations.

Addressing AI Challenges: Identifying informal settlements and climate risks with AI is essential to selecting adaptation and mitigation strategies that can reduce health risks for people living in precarious situations, as well as address health and contextual inequities they face. AI/ML can be further used to predict areas most at risk in urban areas to maximize intervention efficacy and impact.

Links and Publications: Will be presenting some work at the International Conference for Urban Health in November.

Linking Village-Level Data on Climatic Impact-Drivers, Socioeconomic Vulnerability, and Environmental Conditions to health Impacts across 12 Philippine Cities

Climate & Health (2022)

Team Members: Patricia Anne Faustino (Thinking Machines Data Science),* Anica Araneta (Thinking Machines Data Science),*

Primary Institution: Thinking Machines Data Science

Partner Institutions: EpiMetrics, Inc., Manila Observatory, Philippine Action for Community Led Shelter Initiatives

Project Description: The Project Climate Change, Health, and Artificial Intelligence (Project CCHAIN) dataset is a validated, open-sourced linked dataset containing 20 years (2003-2022) of climate,

environmental, socioeconomic, and health dimensions at the barangay (village) level across twelve Philippine cities (Dagupan, Palayan, Navotas, Mandaluyong, Muntinlupa, Legazpi, Iloilo, Mandaue, Tacloban, Zamboanga, Cagayan de Oro, Davao).

Addressing AI Challenges: Our project creates a dataset that we hope can quantify the relationships between climate, environment, social, and health indicators at a high resolution. We envision the dataset to be used for developing disease outbreak prediction models that can inform local action to protect vulnerable communities.

Links and Publications:

- <https://data.humdata.org/dataset/project-cchain>
- <https://adb.eventsair.com/big-data-for-disaster-response-and-management-in-asia-and-the-pacific/agenda>
- [https://www.idmod.org/symposium/current-symposium/;](https://www.idmod.org/symposium/current-symposium/)

Pre-View: Perinatal health outcomes and climate change observatory

Climate & Health (2022)

Team Members: Rodolfo de Carvalho Pacagnella (University of Campinas – Unicamp)*

Primary Institution: University of Campinas - Unicamp

Partner Institutions: Fitec Inovações Tecnológicas

Project Description: Climaterna's data curation seeks comprehensive and varied information fundamental for analyzing the effects of climate change on maternal and perinatal health in Brazil:

1. Demographic Data: Collected by national institutions like the Brazilian Institute of Geography and Statistics (IBGE), including population information, birth and mortality rates, age, sex distribution, and socioeconomic characteristics.
2. Maternal and Perinatal Health Data: Includes information from health registration systems such as hospital information systems, civil birth and death registries, and data collected by the Ministry of Health and public health research institutions.
3. Climate Data: Provided by meteorological and climate research institutions such as the National Institute of Meteorology (INMET) and the National Institute for Space Research (INPE).

How are these data connected?

Integrating demographic, maternal and perinatal health, and climate data is essential to understand and mitigate the impacts of climate change on public health in Brazil. This project provides valuable insights into how climatic factors affect maternal and perinatal health and can guide the formulation of effective policies to protect the most vulnerable populations.

Addressing AI Challenges: The combination of these datasets is unprecedented in Brazil and innovates by connecting this information through five main axes:

1. Integrated Analysis of Demographic and Health Data: cross-referencing maternal and perinatal health information to identify patterns of vulnerability and risk in different population groups in the country, outlining the most affected population profiles by climatic impacts.
2. Correlation between Climate Events and Health Indicators: The relationship between extreme climate events (like heatwaves and floods) and maternal and perinatal health is analyzed to identify correlations, such as the association between exposure to extreme temperatures and increased incidence of preterm births.

3. Predictive Modeling: Models that help anticipate challenges and plan more effective public health interventions in the face of climate risks.
4. Identification of Risk Areas: By combining climate data with geographic and health information, it is possible to identify geographical areas at higher risk of negative climate change impacts on maternal and perinatal health, directing resources and efforts to the most vulnerable areas.
5. Scientific Evidence for Public Policies: This project aims to support the formulation of public policies aimed at mitigating the impacts of climate change on health, with a special focus on vulnerable populations. These policies may include climate adaptation measures and strengthening health infrastructure.

AI is being used for data ingestion, governance, and for the analysis of the interactions.

Tanzania Climate-Sensitive Waterborne Diseases Dataset for Predictive Machine Learning Climate & Health (2022)

Team Members: Joseph Philipo Telemala (Sokoine University of Agriculture)*

Primary Institution: Sokoine University of Agriculture

Partner Institutions: Morogoro Municipal Council Head Office

Project Description: Our dataset includes waterborne disease cases from health facilities across five districts in Tanzania (Morogoro, Dodoma, Singida, Ilala, and Temeke) collected over five years (2018-2023). It also features field data on water sources, toilet facilities, and dumpsites in these districts, alongside open weather data from Copernicus, allowing for a comprehensive AI analysis of environmental and public health interactions that can lead to prediction of waterborne diseases.

Addressing AI Challenges: The data collected from our project can be used to analyze the relationship between environmental factors, such as water sources and waste management, and the prevalence of waterborne diseases. By integrating health and environmental data, we aim to predict disease outbreaks and inform climate-resilient public health strategies in vulnerable communities where availability of data is challenging.

The INSPIRE network: Building a data science platform for integration and harmonisation of longitudinal population health studies and climate data in Africa.

Climate & Health (2022)

Team Members: Agnes Kiragga (African population and Health Research Center)*, Stella Chege

Primary Institution: African population and Health Research Center, Nairobi

Partner Institutions: London School of Hygiene and Tropical Medicine (LSHTM), Makerere Centre for Health and Population Research, Nairobi Urban Demographic and Health Surveillance Site, Malawi Epidemiology and Intervention, National Institute for Medical Research, Mwanza Center (NIMR)

Project Description: The dataset contains remote satellite data and health outcome data including mortality, morbidity from Tanzania and Uganda.

Addressing AI Challenges: The dataset leverages AI platforms to automate data analysis pipelines from health and demographic sites in East Africa.

CLIMATE & FORESTS DATASETS

AI for predicting land use/land cover changes in Uganda

Climate & Forests (2023)

Team Members: Sarah Akello (Makerere University)*

Primary Institution: Makerere University Kamala

Partner Institutions: Initiative for Sustainable Development - Kenya

Project Description: We aim at generating accessible and quality satellite imagery datasets and National Forestry Inventory datasets from Uganda. These datasets will be utilised to predict land cover types or changes for informed decision making in natural resource management. This research will utilise sentinel 2 and tree cover satellite imagery datasets and augment it with the biomass inventory datasets. In partnership with the Uganda National Forestry Authority (NFA), we propose to deliver of a total of 384 sentinel 2 image datasets from 2018 and the hensen tree cover dataset of 384 images. A biomass mean stock per hectare dataset for each vegetation type will also be acquired. Annotation procedure of the different datasets will be further conducted to deliver on each intended machine learning task.

Addressing AI Challenges: Making Predictions from data captured to inform policy, improve practice and add to the body of knowledge.

An integrated dataset combining remote sensing, ground measurements and weather data for improved reforestation monitoring

Climate & Forests (2023)

Team Members: Ciira Maina (Dedan Kimathi University of Technology)*

Primary Institution: Dedan Kimathi University of Technology

Partner Institutions: Data Science Africa, Trans-African Hydro-Meteorological Observatory (TAHMO), Integrated Forestry Consultancy and Management Services Kenya Limited (IFCMS)

Project Description: We are collecting drone data, tree parameter measurements and weather data from an area in Kenya. We hope to monitor growth and survival of the forest stand. In this work we propose to collect a high resolution dataset of drone imagery, ground truth tree parameter measurements and aggregate weather data from a large reforested stand in Kenya. These data will be collected at a high spatial and temporal resolution to enable the development of tools to rapidly assess the success of reforestation efforts by enabling counting of trees and estimation of tree biophysical parameters from drone images and prediction of tree growth in relation to tree species and weather variables. In addition, we will validate the use of a custom built handheld stereo camera for rapid data collection and tree parameter computation. We will release annotated drone imagery with tree crowns and species labels as well as ground based images with tree masks for tree segmentation. We will also make available aggregated data from weather stations around the study area via an API.

Addressing AI Challenges: We hope to use these data to model forest growth and quantify the effect of climate change on reforestation efforts.

Blue carbon and economy in West Africa: mobilizing datasets to better understand mangrove health in relationship to climate change

Climate & Forests (2023)

Team Members: Salako Kolawole Valere (Laboratory of Biomathematics and Forest Estimations, University of Abomey-Calavi, Republic of Benin),* Romain Glèlè, Lucas Kakaï

Primary Institution: Laboratory of Biomathematics and Forest Estimations, University of Abomey-Calavi, Republic of Benin

Project Description: Our project aims to create datasets related to the health of mangrove ecosystems in Benin and their relationship to environmental and climate. Specifically, the datasets will include information on mangrove species distribution, vegetation cover, sediment characteristics, water quality, and carbon stock. In particular, the project will generate 500 annotated drone-based images, 600 observations on water physico-chemical parameters, 1200 observations on soil characteristics, 10800 climate observations, forest inventories data (canopy height and cover, tree diameter, total height, etc.) from 600 plots, and data from 40 Focus Group Discussions and 400 individual semi-structured surveys in two major Ramsar sites.

Addressing AI Challenges: There is a critical gap in real ground data on mangroves in West-Africa and this is hindering the development and testing of AI. For example, how drone images can be used to easily and accurately estimate carbon stock in mangroves? How can available large datasets on climate can be combined with real word field data to understand the potential impact of climate on ecosystems ability to store carbon? Our data will help addressing these pressing questions through by feeding (data supply) the development and testing of AI that address carbon stock in Mangroves, hence climate challenges.

Data-enabled climate shock absorbance through agroforestry (Agrof4resilience)

Climate & Forests (2023)

Team Members: David Makori (International Centre of Insect Physiology and Ecology),* Bester Mudereri

Primary Institution: International Centre of Insect Physiology and Ecology (icipe)

Partner Institutions: Kenya Forestry Research Institute (KEFRI), Strathmore University

Project Description: Agroforestry data in Kenya is collected to explore the effect of climate change on agroforestry as a climate-resilient, gender-inclusive and climate-smart source of income. Precise locations and configurations of these Agroforestry Systems (AFSs) are relatively unknown in Kenya. This makes AFSs' carbon market evaluation, spatial sampling, and mapping exercises relatively challenging. We propose to collect a huge AFSs dataset, sustainably maintain an open-access dataset and provide documentation to guide the public on how to access and use it. These datasets will be machine-learning ready for further analytics and modeling using machine-learning algorithms for purposes of AFS optimization, species identification, and sequestration valuation. A minimum of 20,000 land use points will be collected in four transects across the agroecological zones of Kenya (35 out of 47 counties), while 200 ha will be sampled for tree species identification, height, and diameter at breast height (DBH) with a total of 15,000 individual plants sampled, photographed, and geotagged.

Addressing AI Challenges: We will study a lone effect of climate change on agroforestry using AI algorithms.

Digital MRV for Himalayas

Climate & Forests (2023)

Team Members: Michael Anthony (Verify.Earth),* Jo Dawson

Primary Institution: Verify.Earth

Partner Institutions: Balipara Foundation, Earth Analytics India Pvt. Ltd.

Project Description: Forest degradation is wide-spread in the Himalayas. Causes are diverse but often linked to deteriorating economic livelihoods. Balipara Foundation (India) is a leading conservation organisation in India, having reversed forest degradation in North-eastern India across an estimated 5000 ha through creating alternative livelihoods for indigenous communities. While Balipara is collecting ground data on forest biomass, degradation incidences and socioeconomic factors, it finds it difficult to apply that data spatially so that it can be used to map potential areas for new restoration interventions and to consistently monitor the nature-impact of its work - potentially also for registering forest carbon credits. To that aim, Earth Analytics India Pvt Ltd will develop an algorithm that is making use of new radar satellite sensors to overcome the shortcomings of existing approaches, notably their rare temporal frequency. The ground data collection approach will be led by youth members of Indigenous Hubs that Balipara Foundation has built in two states in the Himalayas. They will be guided by the science team of Balipara in collaboration with local universities and will focus on biodiversity and forest parameters to be used by the geospatial data team for calibration.

Addressing AI Challenges: We want to understand the drivers and dynamics of deforestation better and also correlate them to living incomes of indigenous communities in the area that we are working in.

Forest Carbon Sequestration in the Congo Basin

Climate & Forests (2023)

Team Members: Kenmoe Kendie (World Resources Institute)*, Nancy Harris

Primary Institution: World Resources Institute

Partner Institutions: RECOFTC, Purdue University, University of Yaounde, University of Kisangani

Project Description: We will to the grantees Maps of historical and projected changes in forest carbon stocks under current and future climate change scenarios for Congo Basin countries. This project will re-measure 65 field-sampled plots across forests of the Democratic Republic of Congo (DRC) and Cameroon. Data collected from these plots will be merged into a database of more than 1.5 million plots worldwide and used to train and validate a geospatial AI-based forest growth model. Maps produced from this project will show past, current and future dynamics of tree species distribution, species dominance and associated changes in forest carbon sequestration, biodiversity and timber quality over time. The data can be used to set credible and locally relevant forest reference levels that can unlock the 'plus' of available REDD+ carbon finance for the sustainable management of forests and for the conservation and enhancement of forest carbon stocks

Addressing AI Challenges: Data collected under first part of the projet will be integred into the re-measured global forest inventory database (GFI-4D). This will be used to train and validate an AI-based stand-level forest growth model (MATRIX) that combines plot re-measurements with remote sensing observations and other geospatial data layers to produce raster output maps at 1-km spatial resolution (or finer) of historical and projected future changes in forest carbon stocks over time back to the year 2000 and forward to the year 2070. Additional information will also be produced on projected changes to tree species composition over time.

Machine learning training data for continental Ecuador and Galapagos

Climate & Forests (2023)

Team Members: Cristina Aguilar (Fundación EcoCiencia),* Carmen Josse, Maria Olga Borja

Primary Institution: Fundación EcoCiencia

Partner Institutions: Armed Forces Politechnical School (ESPE University)

Project Description: A georeferenced dataset composed of 30,000 points that will be distributed across continental Ecuador and the Galapagos islands (an area equivalent to 257,215.3 Km²).

Each point will have been precisely inspected by trained annotators and will contain attributes on:

1) land use/cover category (out of 20 -natural and artificial- possible categories) for each year between 1985 and 2022; and 2) conservation status (presence/absence within protected area and/or indigenous territory).

Addressing AI Challenges: Land cover/use temporal trends are essential to understand the status of the remaining natural cover and understanding climate change trends. When combined with status of protection, land cover/use data can reveal historical trends on what is most effective as societal actions to assure the existence of our remaining natural capital and the ecosystem services that are thus provided by them. Using precisely inspected georeferences dataset on AI models can facilitate the assessment of ongoing and future trends of change in remaining ecosystems and their rate of change, which have the potential to guide adjusted action for policymakers and investors.

LTOME-KATIP Project: Indigenous-Led Labelling for Inclusive AI in Addressing Human-Wildlife Conflict and Climate Change

Climate & Forests

Team Members: Diana Mastracci (GEO Indigenous Alliance)*

Primary Institution: GEO Indigenous Alliance

Project Description: The Ltome-Katip Indigenous-Led Monitoring for Human-Wildlife Conflict Mitigation and Climate Resilience project is a pioneering initiative dedicated to generating Indigenous-labelled data for machine learning. Spearheaded by the Shuar tribe in the Ecuadorian Amazon and the Samburu tribe in northern Kenya, in collaboration with the Rochester Institute of Technology, Space4Innovation, and the GEO Indigenous Alliance this project addresses the urgent need for Indigenous-labelled data in machine learning. Focusing on regions heavily impacted by climate change, namely the Amazon jungle and northern Kenya, our mission is to revolutionize approaches to human-wildlife conflict and climate resilience. By actively involving Indigenous communities in data collection, labelling, and map development, we bridge critical gaps in understanding climate change, human-wildlife conflict, and afforestation/deforestation. This collaboration not only safeguards invaluable biodiversity and preserves Indigenous identity but also empowers communities, ensuring culturally relevant and effective solutions. Through the integration of Indigenous knowledge and cutting-edge technology, we are poised to make a tangible difference in the lives of both people and wildlife in these vital regions while reducing bias and uncertainty in AI models.

Quantifying Colombian Mangroves Aboveground Biomass and Carbon Content

Climate & Forests (2023)

Team Members: Isabela Katime Arroyave*, Maria Cuevas-Gonzalez

Primary Institution: Centre Tecnològic de Telecomunicacions de Catalunya (CTTC)

Project Description: Accurate information on AboveGround Biomass (AGB) is crucial for various purposes such as improving carbon accounting methods, guiding conservation initiatives, monitoring climate change impacts and mitigation efforts, and supporting sustainable management practices. This

proposal aims to address the need for reliable and up-to-date information on AGB and AboveGround Carbon (AGC) in the mangrove ecosystem of the Colombian Caribbean. By creating and compiling a comprehensive dataset, the project aims to provide inputs for machine and deep learning applications that can accurately estimate AGB and AGC. The dataset will consist of field data collected through on-site measurements, UAV surveys and satellite-derived data. Field data and UAVs will provide ground-truth information, while satellite data will offer broader coverage and enable the assessment of AGB/AGC at a larger scale. Machine learning applications can leverage this dataset as inputs for training and validating models to accurately estimate AGB. By employing advanced algorithms, these models can analyze and exploit the relationships between field and remotely sensed data to develop robust predictions of AGB/AGC across mangrove ecosystems. The proposed dataset will serve as a valuable resource for researchers, conservation practitioners, policymakers, and other stakeholders involved in mangrove ecosystem management. By addressing the need for reliable and up-to-date AGB/AGC information, the project aims to enhance understanding and decision-making related to mangrove conservation, carbon sequestration, and sustainable development in the Colombian Caribbean. The methodology developed through this project might even be applied beyond the Colombian Caribbean to other mangrove ecosystems with similar structures.

Reference Data Collection for Improving Land Use Change Mapping in Ghana

Climate & Forests (2023)

Team Members: Foster Mensah (Center for Remote Sensing and Geographic Information Services),* Jacob Abramowitz

Primary Institution: Center for Remote Sensing and Geographic Information Services (CERSGIS)

Partner Institutions: The University of Alabama in Huntsville, Boston University, International Crops Research Institute for the Semi-Arid Tropics (ICRISAT)

Project Description: Roughly 20,000 cocoa field polygons and 20,000 other land cover points of interest (small-scale mining, oil palm, rubber, degraded/open forest, other agriculture) are provided as point shapefiles. Additionally, it is expected that roughly 2,000 socioeconomic surveys will be completed.

Addressing AI Challenges: The reference data collected by our project will fuel AI/machine learning-based land use change monitoring applications in Ghana that are otherwise not possible

Southern African Trees for Climate Resilience: A Comprehensive Bioinformatics Approach

Climate & Forests (2023)

Team Members: Heath Beckett (School for Climate Studies, Stellenbosch University),* Guy Franklin Midgley, Julia Harper

Primary Institution: Stellenbosch University School for Climate Studies

Partner Institutions: Phyla.Earth, Weihenstephan Triesdorf University of Applied Sciences

Project Description: Our dataset comprises comprehensive floristic and plant functional information for Southern African trees, paired with biogeoclimatic data. This integrated database will enable us to model the ecological success and growth patterns of these trees under current and projected future climate conditions. The dataset will serve as a foundation for assessing climate risks to Southern African flora, identifying resilient indigenous species, mapping ideal locations for conservation efforts, and quantifying the carbon mitigation potential of these trees across various regional conditions.

Addressing AI Challenges: Our project leverages AI to analyze complex biogeoclimatic data and tree characteristics, enabling accurate predictions of how Southern African trees will respond to climate change. By using machine learning algorithms, we can identify the most resilient indigenous species and optimal locations for reforestation efforts, directly addressing climate adaptation challenges. Additionally, AI-driven analysis of our dataset will help quantify the carbon sequestration potential of various tree species, contributing to targeted and effective climate mitigation strategies in the region.

WatchMyTree

Climate & Forests (2023)

Team Members: Fabrice Zapfack (data354),* Therence Temfak, Grégoire Kania

Primary Institution: data354

Partner Institutions: Actum Dev, Ministry of Environment and Sustainable Development, Ministry of Environment and Sustainable Development

Project Description: The Côte d'Ivoire Aboveground Biomass (AGB) dataset is a comprehensive collection of biomass measurements from 263 plots across various regions of Côte d'Ivoire, collected between 2021 and 2023. This dataset, essential for sustainable forest management and carbon stock estimation, includes key parameters such as tree height, diameter at breast height, density, and scientific name, with the final biomass calculated using allometric equations. It serves as a foundational resource for developing AI models in the African context and supports research and innovation in forest resource management.

Addressing AI Challenges: By enabling precise estimation of forest aboveground biomass (AGB), which is crucial for monitoring carbon stocks and managing forests sustainably. By leveraging AI models trained on the Côte d'Ivoire AGB dataset, the project enhances our ability to predict forest growth trends and assess the impact of climate change, contributing to informed decision-making and effective conservation strategies.