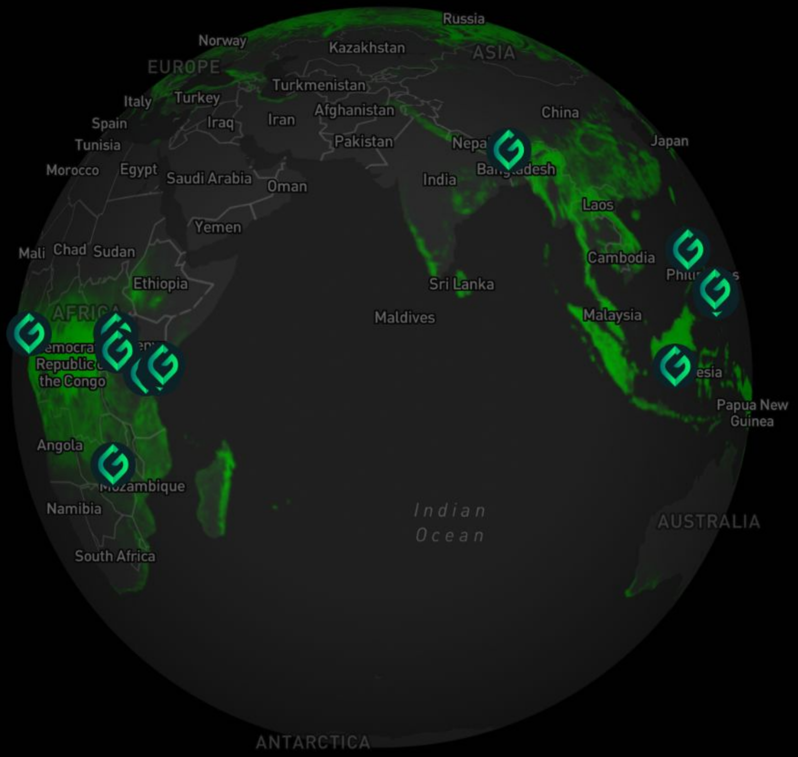
Forests & Land Use

CCAI Summer School



David Dao

@dwddao

About me


CELO





BIENVENIDOS
PARQUE NACIONAL
DEFENSORES DEL CHA
MADES

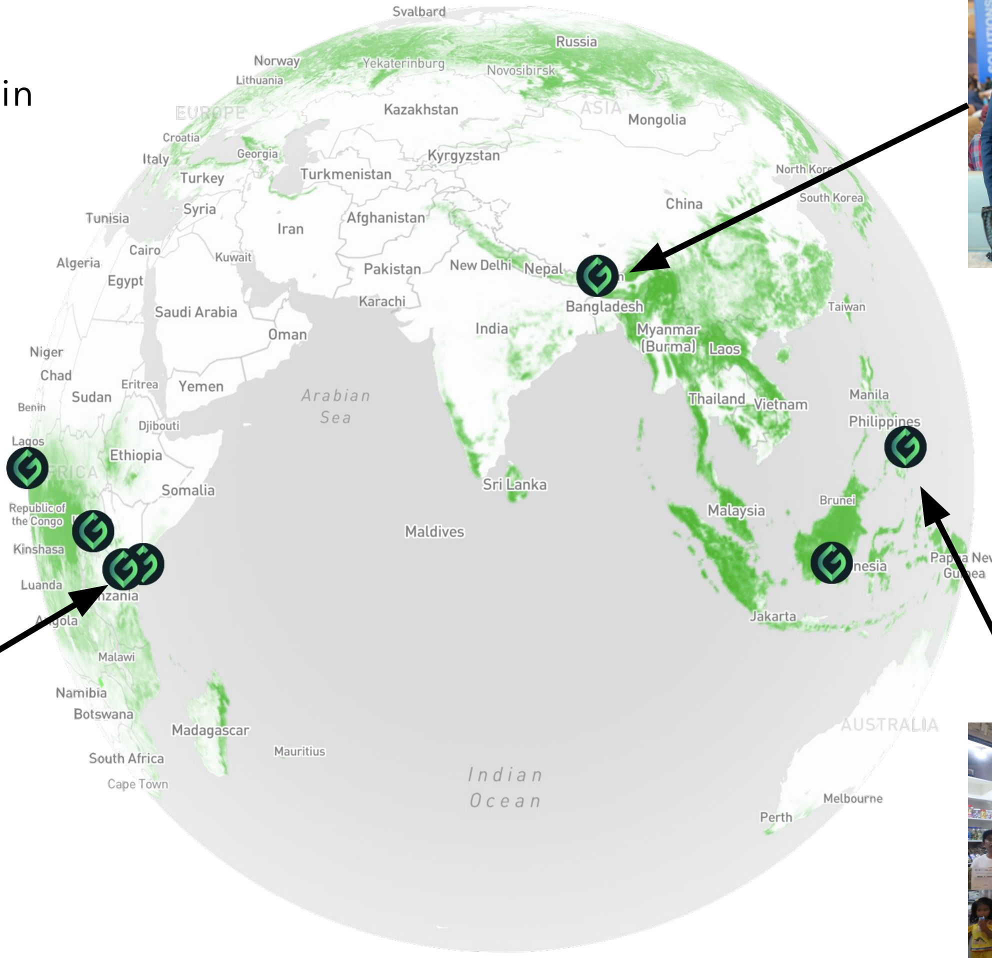
TIME FOR ACTION
OP25
HILE
ID 2019
HANGE CONFERENCE

MADES

About GainForest

27 Partner Organisations in the Global South

Kenya



Bhutan

Philippines

Lecture Overview

1. Climate Change: Forests and Land Use
2. Financing for Nature: Carbon Credits and PES
3. Artificial Intelligence: How can it help?
4. Data Exploration with GainForest (**Practice**)
5. Lessons learned for AI for Forests (**Discussion**)

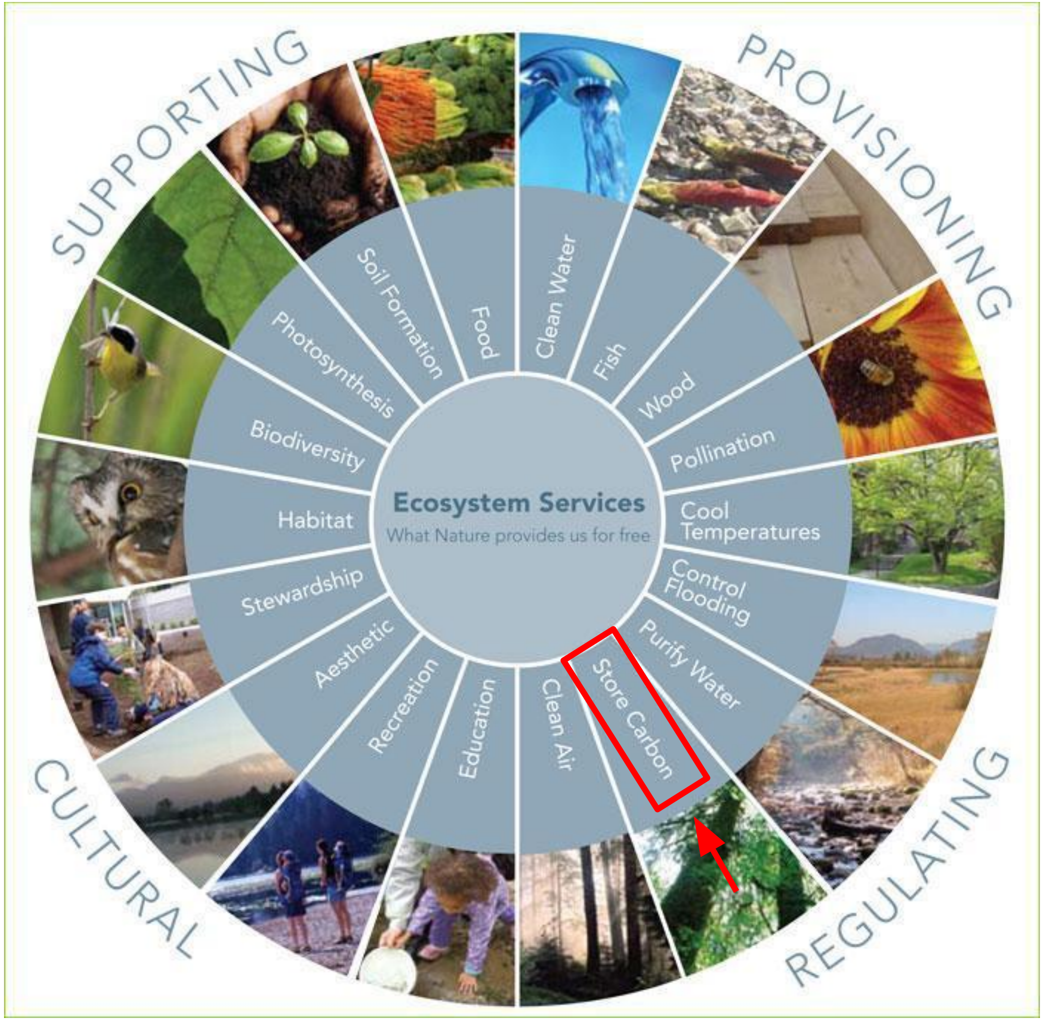
I. Climate Change

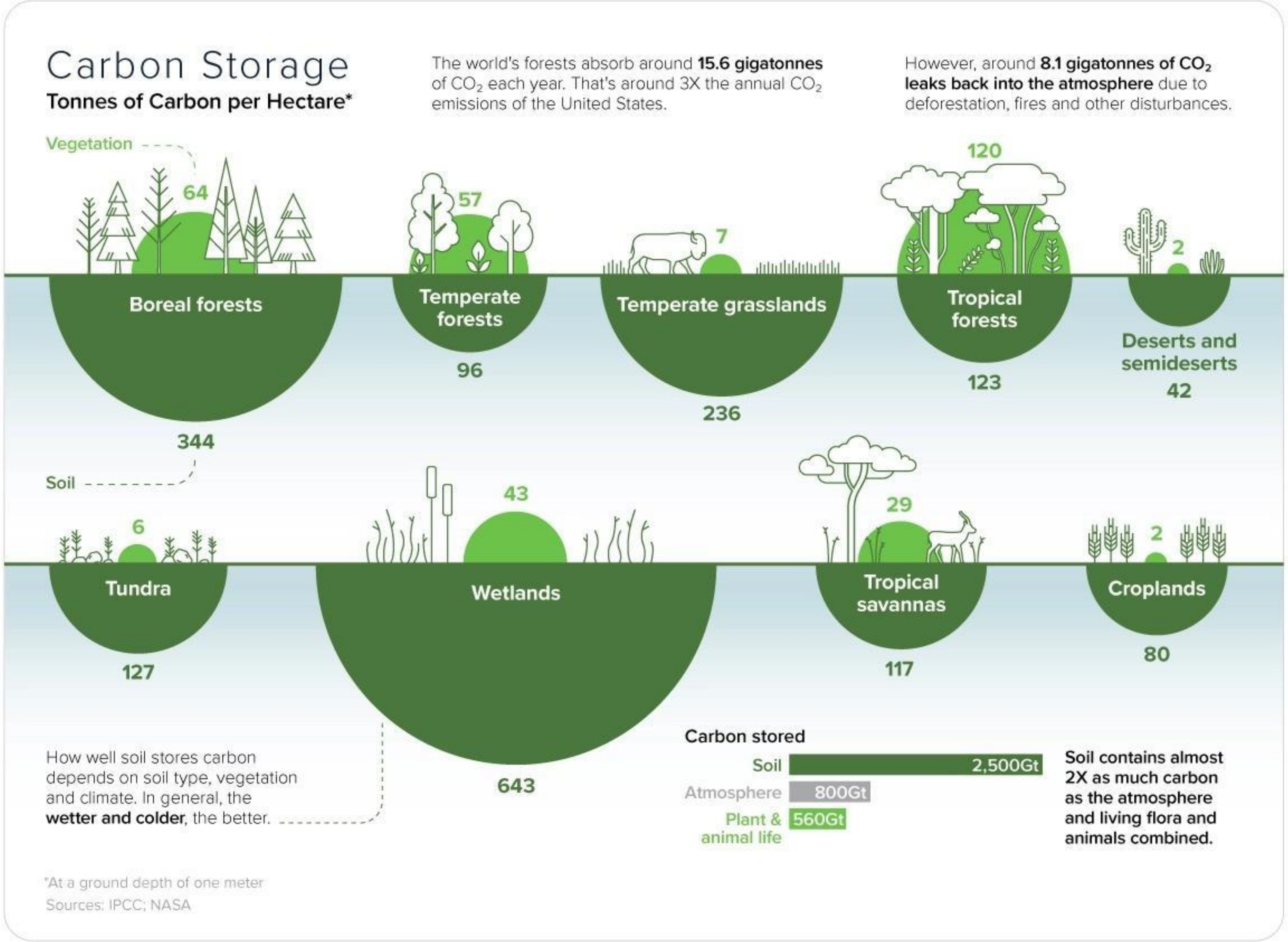
Forests and Land Use

Listen 🌱

“Nature is essential for human existence and good quality of life. Most of nature’s contributions to people are not fully replaceable, and some are irreplaceable.”

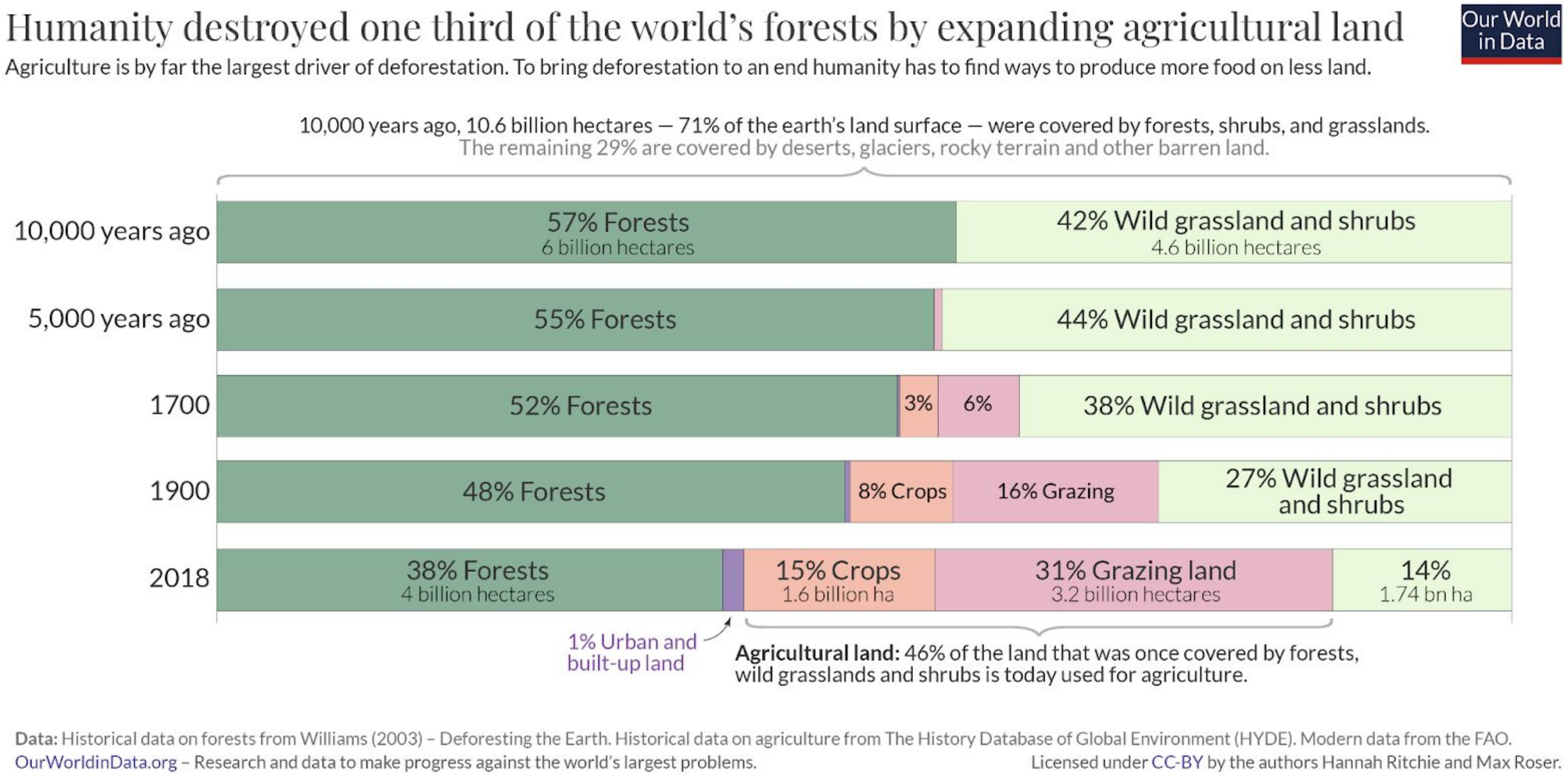
The Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services


SUPPORTING
PROVISIONING
REGULATING
CULTURAL
Ecosystem Services
What Nature provides us for free
Food
Clean Water
Fish
Wood
Pollination
Cool Temperatures
Control Flooding
Purify Water
Store Carbon
Clean Air
Education
Recreation
Aesthetic
Stewardship
Habitat
Biodiversity
Photosynthesis
Soil Formation



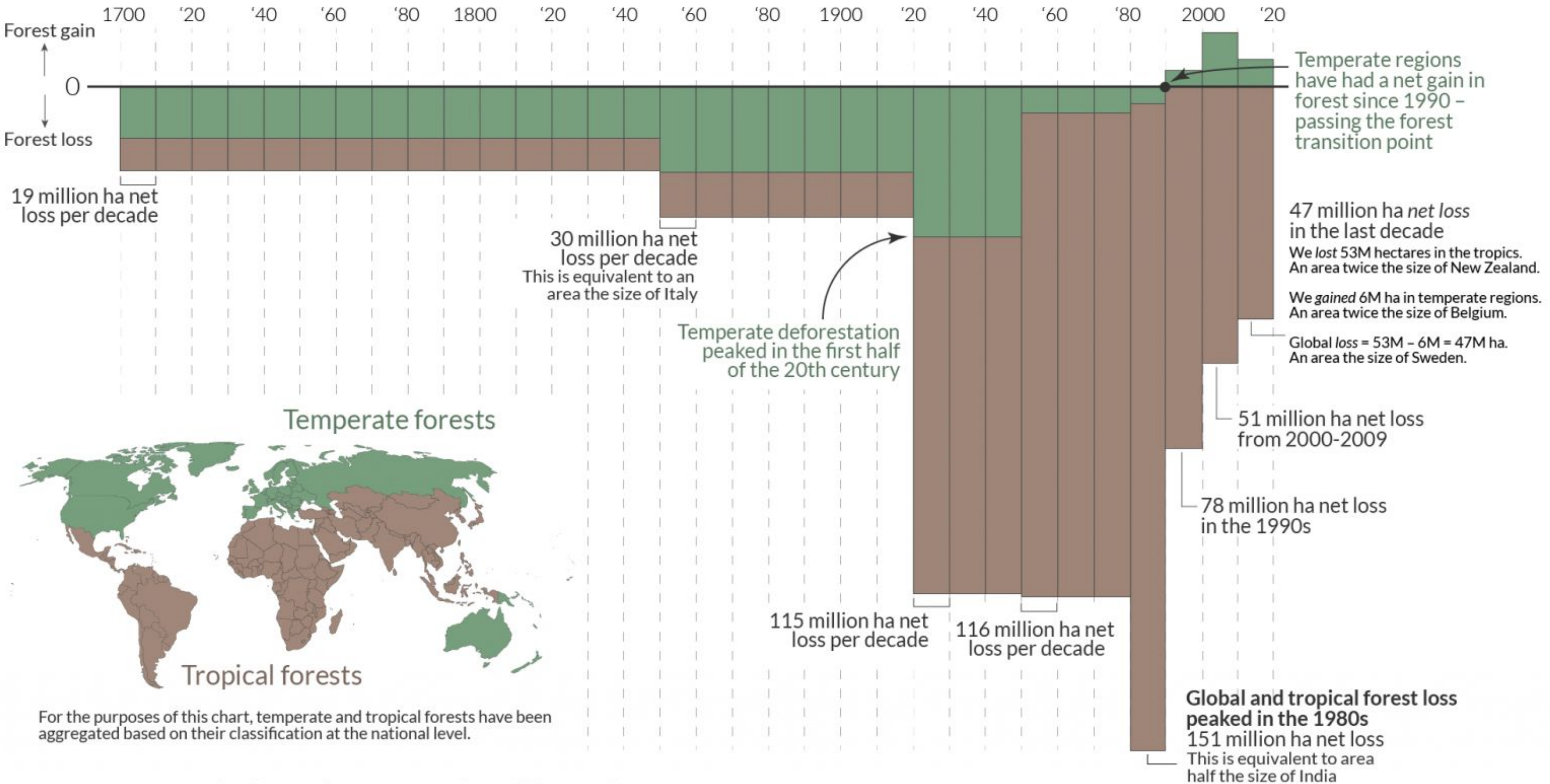
Carbon Storage
Tonnes of Carbon per Hectare*
The world's forests absorb around **15.6 gigatonnes** of CO_2 each year. That's around 3X the annual CO_2 emissions of the United States.
However, around **8.1 gigatonnes of CO_2 leaks back into the atmosphere** due to deforestation, fires and other disturbances.
Vegetation
Soil
Boreal forests
64
344
Temperate forests
57
96
Temperate grasslands
7
236
Tropical forests
120
123
Deserts and semideserts
2
42
Tundra
6
127
Wetlands
43
643
Tropical savannas
29
117
Croplands
2
80
How well soil stores carbon depends on soil type, vegetation and climate. In general, the **wetter and colder**, the better.
Carbon stored
Soil 2,500Gt
Atmosphere 800Gt
Plant & animal life 560Gt
Soil contains almost 2X as much carbon as the atmosphere and living flora and animals combined.
*At a ground depth of one meter
Sources: IPCC; NASA


Location
Mato Grosso, Brazil
1984
“The natural world is deteriorating in rates unparalleled in human history”
The Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services
Google Earth


Humanity destroyed one third of the world's forests by expanding agricultural land
Our World in Data
Agriculture is by far the largest driver of deforestation. To bring deforestation to an end humanity has to find ways to produce more food on less land.
10,000 years ago, 10.6 billion hectares — 71% of the earth's land surface — were covered by forests, shrubs, and grasslands.
The remaining 29% are covered by deserts, glaciers, rocky terrain and other barren land.
10,000 years ago
57% Forests
6 billion hectares
42% Wild grassland and shrubs
4.6 billion hectares
5,000 years ago
55% Forests
44% Wild grassland and shrubs
1700
52% Forests
3%
6%
38% Wild grassland and shrubs
1900
48% Forests
8% Crops
16% Grazing
27% Wild grassland and shrubs
2018
38% Forests
4 billion hectares
15% Crops
1.6 billion ha
31% Grazing land
3.2 billion hectares
14%
1.74 bn ha
1% Urban and built-up land
Agricultural land: 46% of the land that was once covered by forests, wild grasslands and shrubs is today used for agriculture.
Data: Historical data on forests from Williams (2003) – Deforesting the Earth. Historical data on agriculture from The History Database of Global Environment (HYDE). Modern data from the FAO.
OurWorldinData.org – Research and data to make progress against the world's largest problems.
Licensed under CC-BY by the authors Hannah Ritchie and Max Roser.

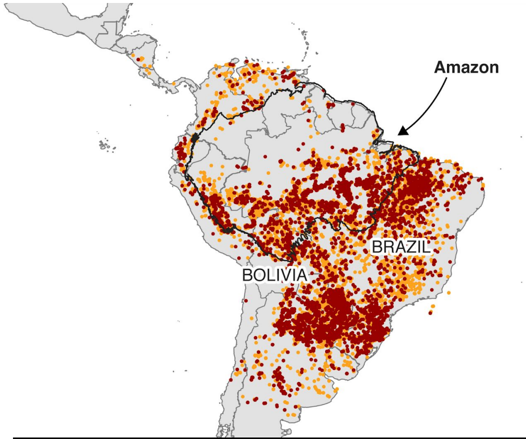
Decadal losses in global forest over the last three centuries

Decadal forest loss is measured as the average net loss of forest area every ten years, in hectares.
This equals deforestation minus any increases in forest area through afforestation.
1.5 billion hectares of global forest was lost between 1700 and 2020 – this is equal to an area 1.5-times the size of the USA.



OurWorldinData.org – Research and data to make progress against the world's largest problems.
Data sources: Based on pre-1990 data from Williams (2006). Deforesting the Earth. Post-1990 data from UN FAO Global Forest Resources Assessment (2020).

Deforestation is a key driver of the climate crisis



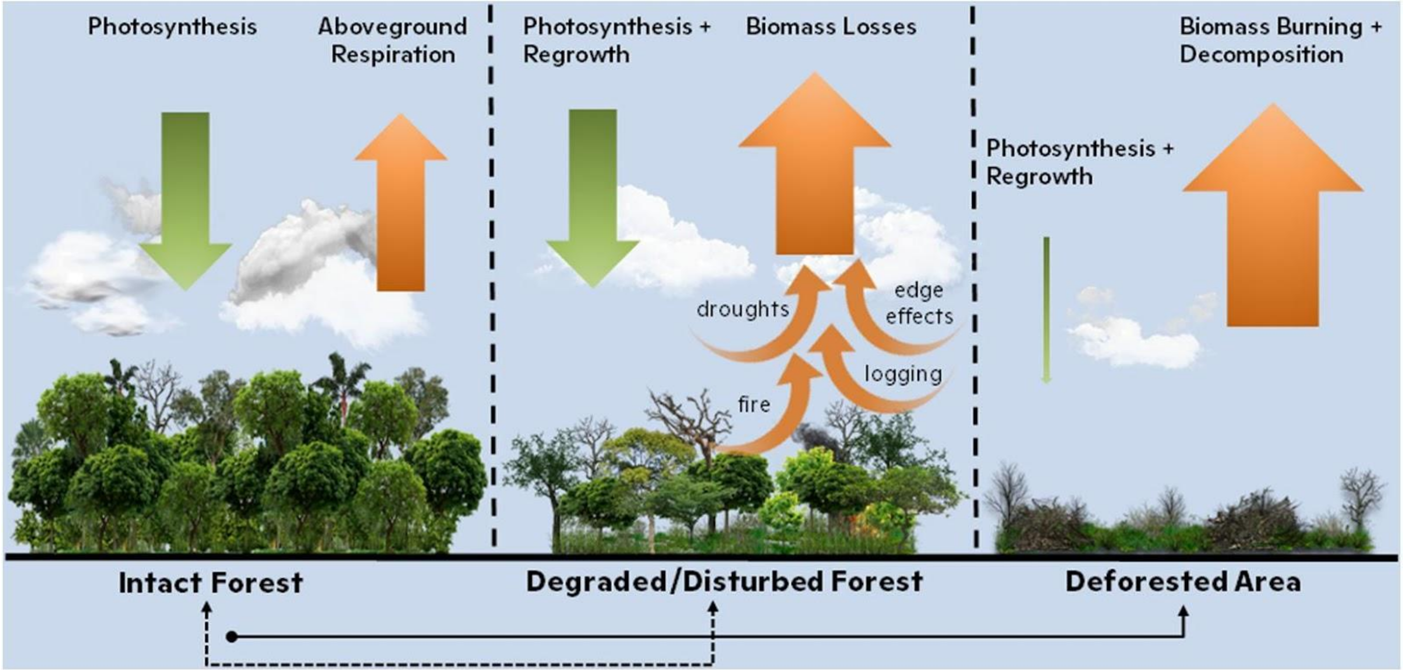
Through deforestation, the Amazon rainforest turned from a carbon sink into a carbon source

Source: MODIS, PlanetLabs imagery

18%

Global anthropogenic emissions

Deforestation and forest degradation



Beyond Deforestation: Carbon Emissions From Land Grabbing and Forest Degradation in the Brazilian Amazon, *Kruid et al.* Frontiers
https://doi.org/10.3389/ffgc.2021.645282

Drivers of Forest Loss (Quiz ❓)

a)

b)

c)

d)

e)

Drivers of Forest Loss (Quiz❓)

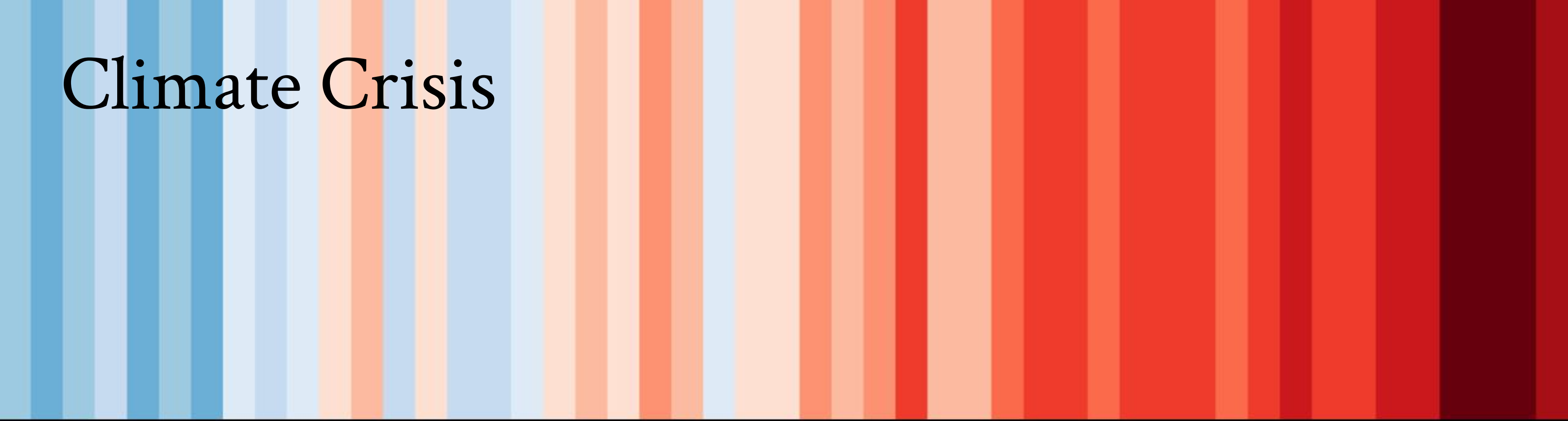
Deforestation
(27%)

Urbanization
(1%)

Shifting
agriculture
(24%)

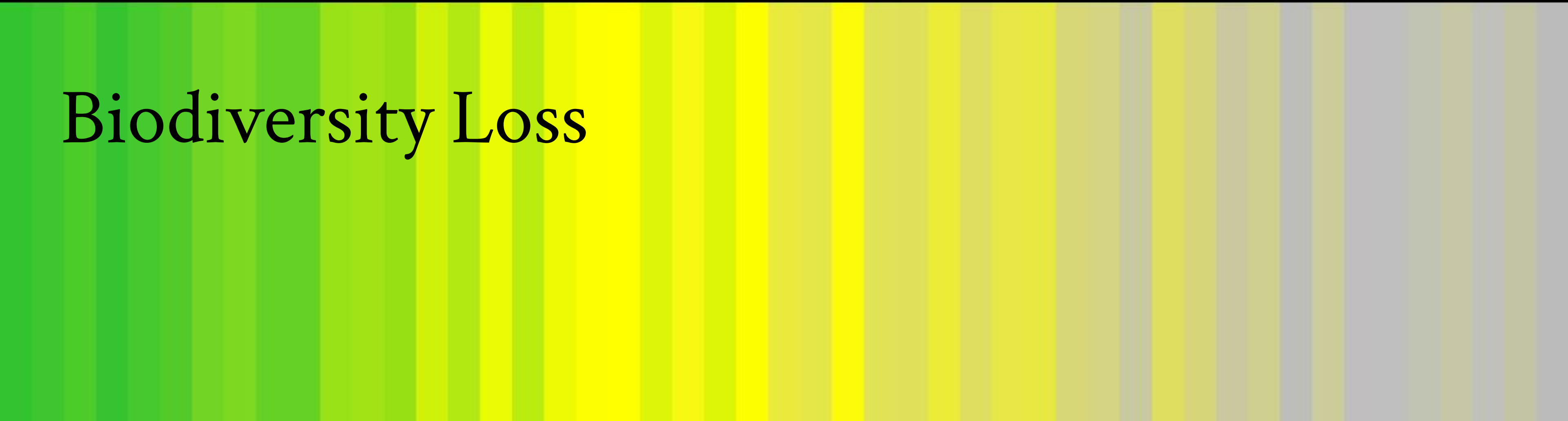
Forest
Products
(26%)

Wildfires
(23%)



Global warming and biodiversity loss 1970 – 2018

Global Warming Stripes #showyourstripes. Data Source UK Met Office CC BY 4.0
From biodiversitystripes.info Data: LPI 2022. Living Planet Index http://stats.livingplanetindex.org/



“Deforestation is changing our climate, harming people and the natural world. We must, and can, reverse this trend.”

Jane Goodall

One Trillion Trees Reforestation Potential

10%

Of emission cuts needed to stay below 1.5 deg*

Global Reforestation Potential, Science 2019, Bastin *et al.*

*rough estimate (but *we* all agree that restoration done right has incredible benefits)

II. Financing for Nature

Carbon Credits and Payments for Ecosystems

Forest Carbon Credits

Reforestation

Planting young trees to remove carbon

Estimated Impact: 15 tonnes of CO_2 per hectare per year

Improved Forest Management (IFM)

Preserve middle-aged trees (captures and stores carbon)

Avoided Deforestation

Preserve old-growth forest

Estimated Impact: 400 tonnes of CO_2 per hectare *(average Amazon rainforest)*

Alternative Financing for Nature

Payment for Ecosystem Services (PES)

Payment for Ecosystem Services (PES) is a concept where landowners are compensated for managing their land in a way that provides some sort of ecological service.

Challenge: Land rights, finding the right KPIs, monitoring

Agroforestry

Agroforestry is a land use management system in which trees or shrubs are grown around or among crops or pastureland.

Challenge: Market access and extensive domain expertise

Ecotourism

Ecotourism is a form of tourism that is focused on the conservation of nature and the well-being of local people.

Challenge: No overuse and equitable share of benefits to communities

Monitoring, Reporting and Verification (MRV)

or “We cannot value what we cannot measure”

Third-Party Monitoring, Reporting and Verification (MRV) is manual and expensive

Tree metrics e.g. Diameter at Breast Height (DBH), species, etc

Allometric equations

Aboveground biomass (AGB) kg

Biomass = 2x Carbon

Carbon stock kg of CO_2

300 USD/Ha

for forest monitoring, verification and reporting

Source: FAO

Results-based payments require monitoring, reporting and verification (MRV)

$25bn

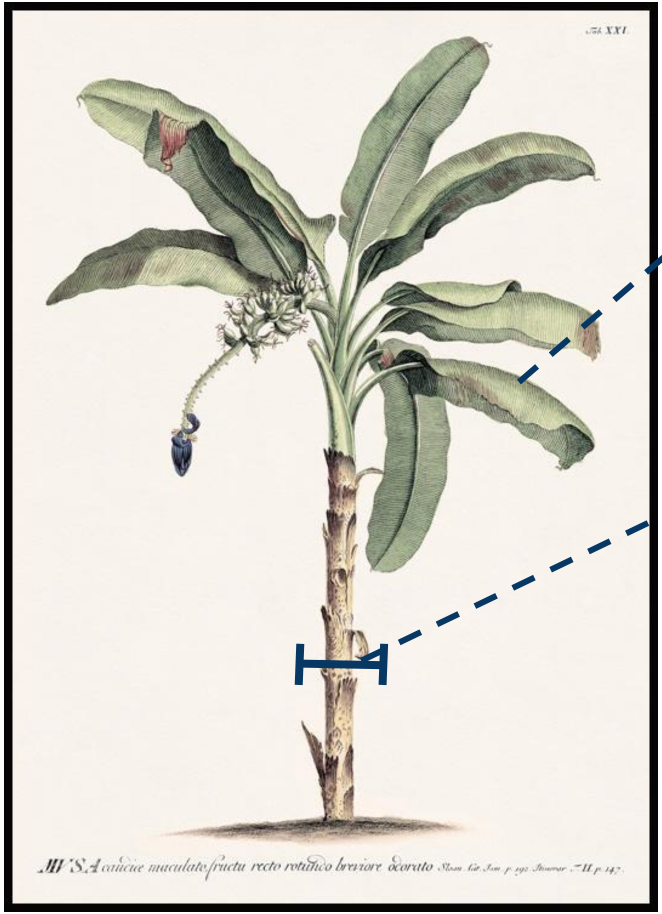
Public funds reserved for results-based payments in forestry

But overlapping land claims, slow performance measurement and missing trust

Make funds inaccessible to local communities

Source: REDD+

Verification?



Species: **Musaceae**

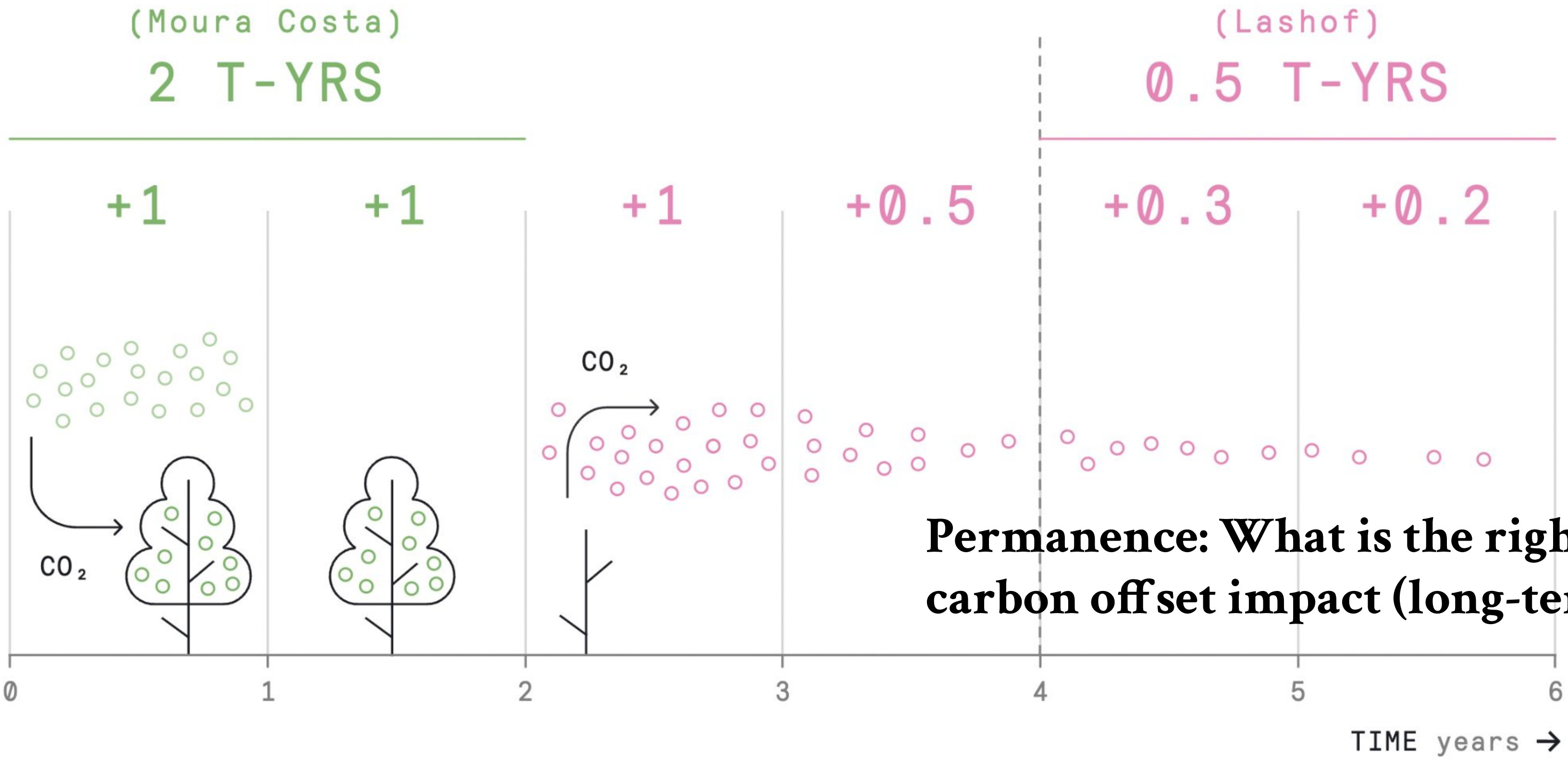
$$AGB_{musacea} = 0.030 * DBH^{2.13}$$

Diameter at Breast Height (DBH) : **0.21 m**

$$AGB_{musacea} = 0.030 * \boxed{0.21}^{2.13}$$

Biomass: **0.001 kg/m2**

Permanence?

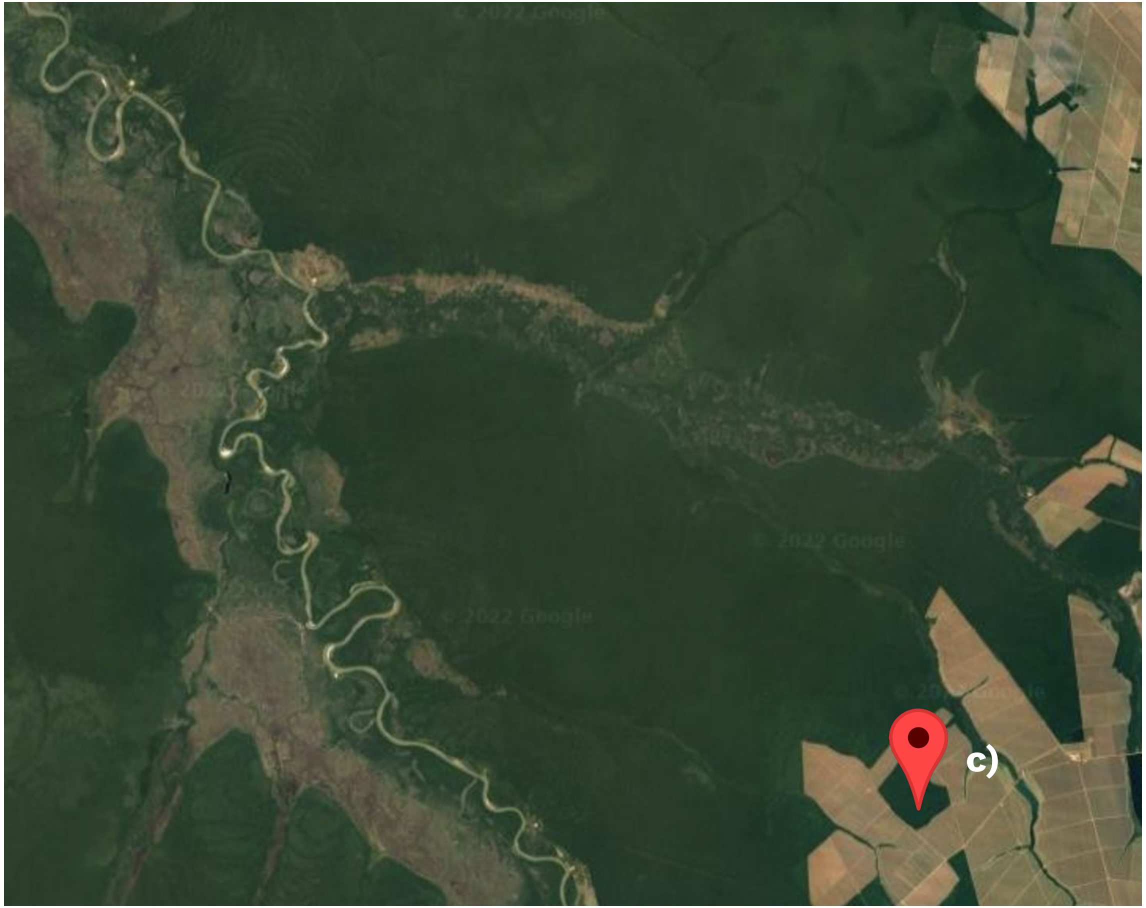


Permanence: What is the right carbon offset impact (long-term)?

Additionality? (Quiz❓)

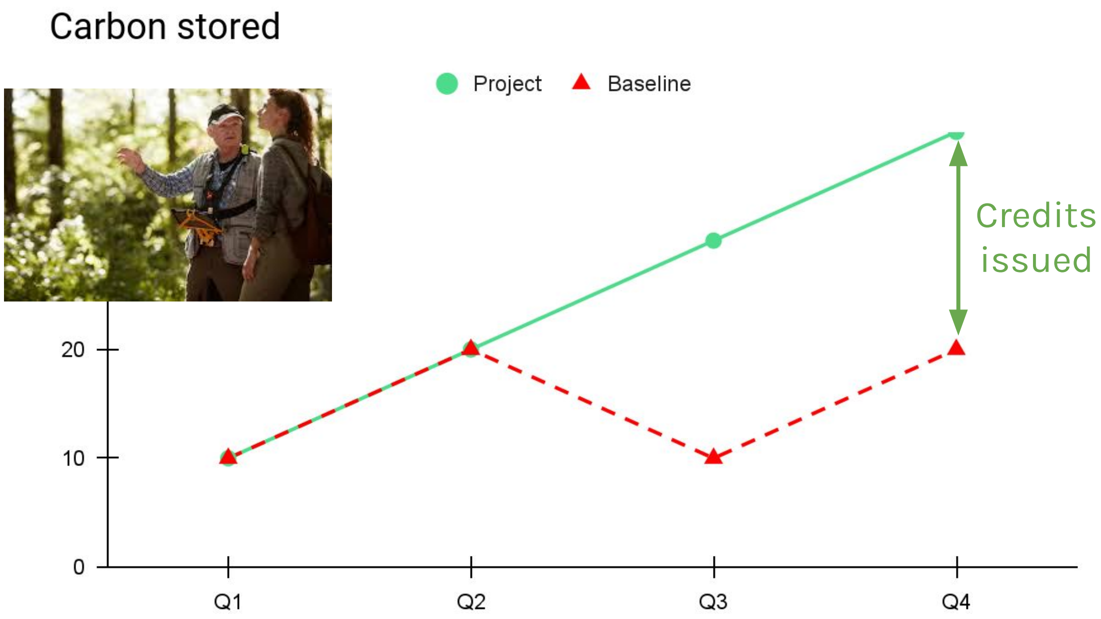
Additionality: What was the risk of deforestation in the first place?

Leakage?



Leakage: Maybe c) was protected but owner just cashed in and instead deforested surrounding

Baseline?



<u>Baseline</u>: What is the right baseline to issue credits? What would have happened without the funding?

Is it greenwashing?

CLIMATE HOME NEWS

Home | News | Comment | Sponsored | Newsletters

Politics | Finance | Justice | Energy | Land | Transport | Science

Five years after New York declaration, forest promises go unmet

Published on 12/09/2019, 4:15pm

Governments and businesses are not living up to voluntary commitments to halve tropical forest loss and restore 150 million hectares by 2020, report finds

This article is more than **3 months old**

Revealed: more than 90% of rainforest carbon offsets by biggest certifier are worthless, analysis shows

Investigation into Verra carbon standard finds most are 'phantom credits' and may worsen global heating

Listen again 😰

Co-Benefits?

Important Note

Carbon Credits

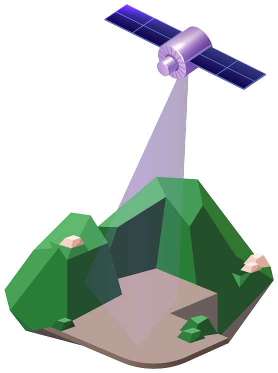
without safeguards can harm climate and environment

Artificial Intelligence

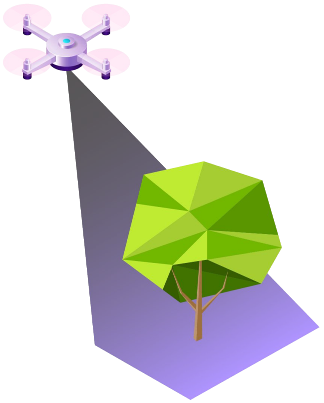
How can it help?

Explosion of multi-modal environmental data

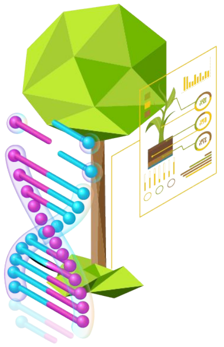
Field-based monitoring



Satellite-based monitoring

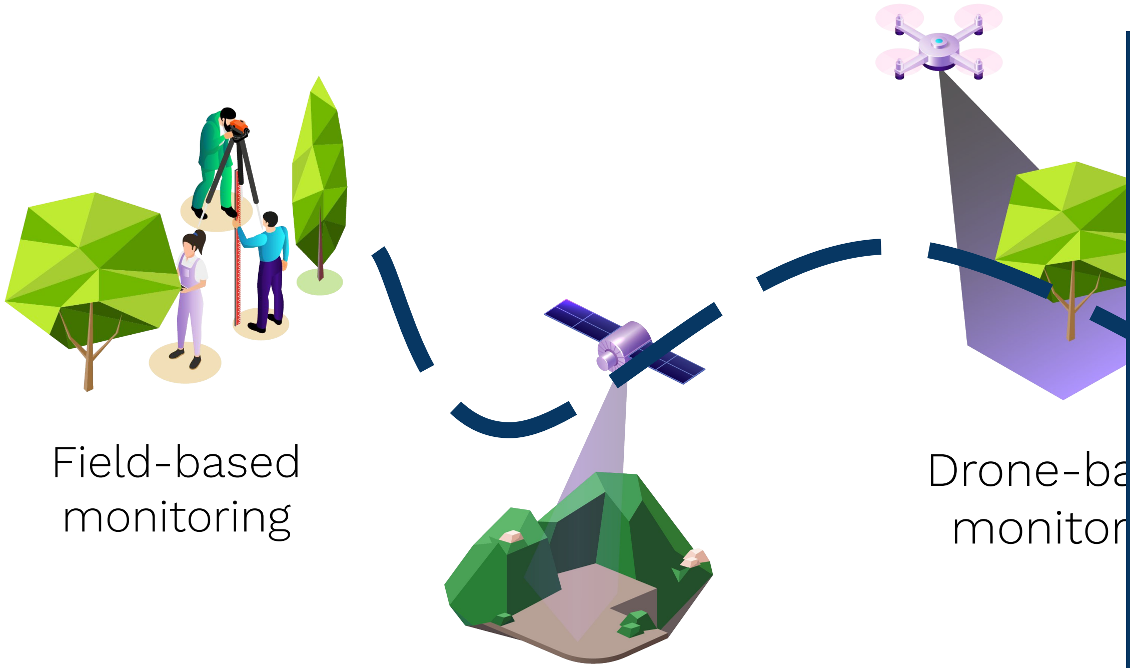


Drone-based monitoring



eDNA monitoring

How can we leverage this data?

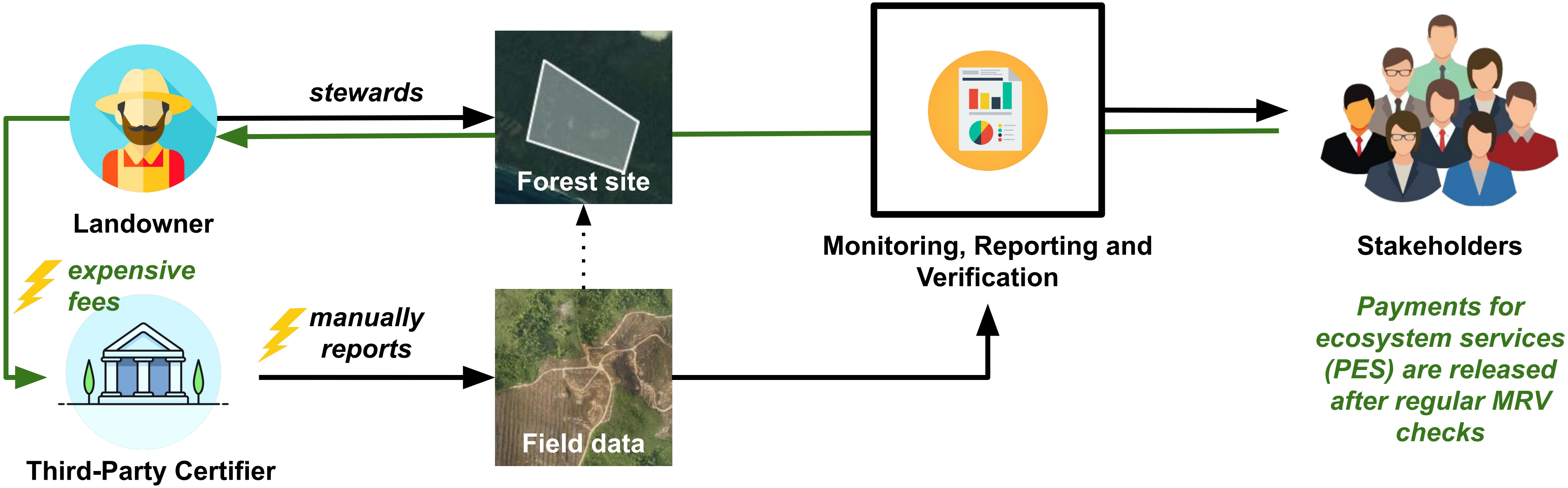


Fusing and learning from multi-modal environmental data provide us with novel *opportunities* (and *challenges*)

AI is uniquely positioned to do so!

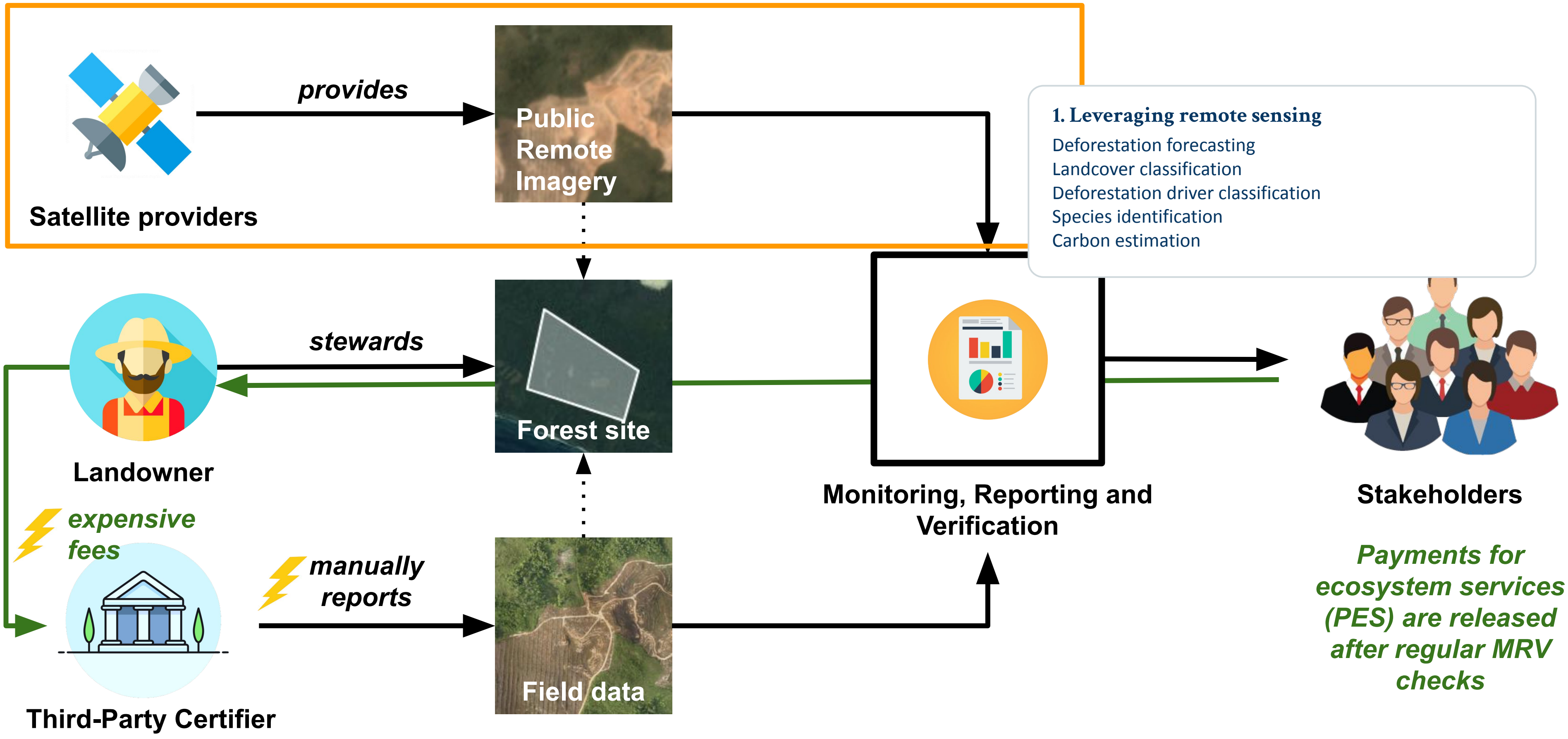
“ML can help us *understand the location, health and ecological value of nature and biodiversity at scale*, and ensure these metrics are reflected in policy, finance, and decision-making.”

Overview



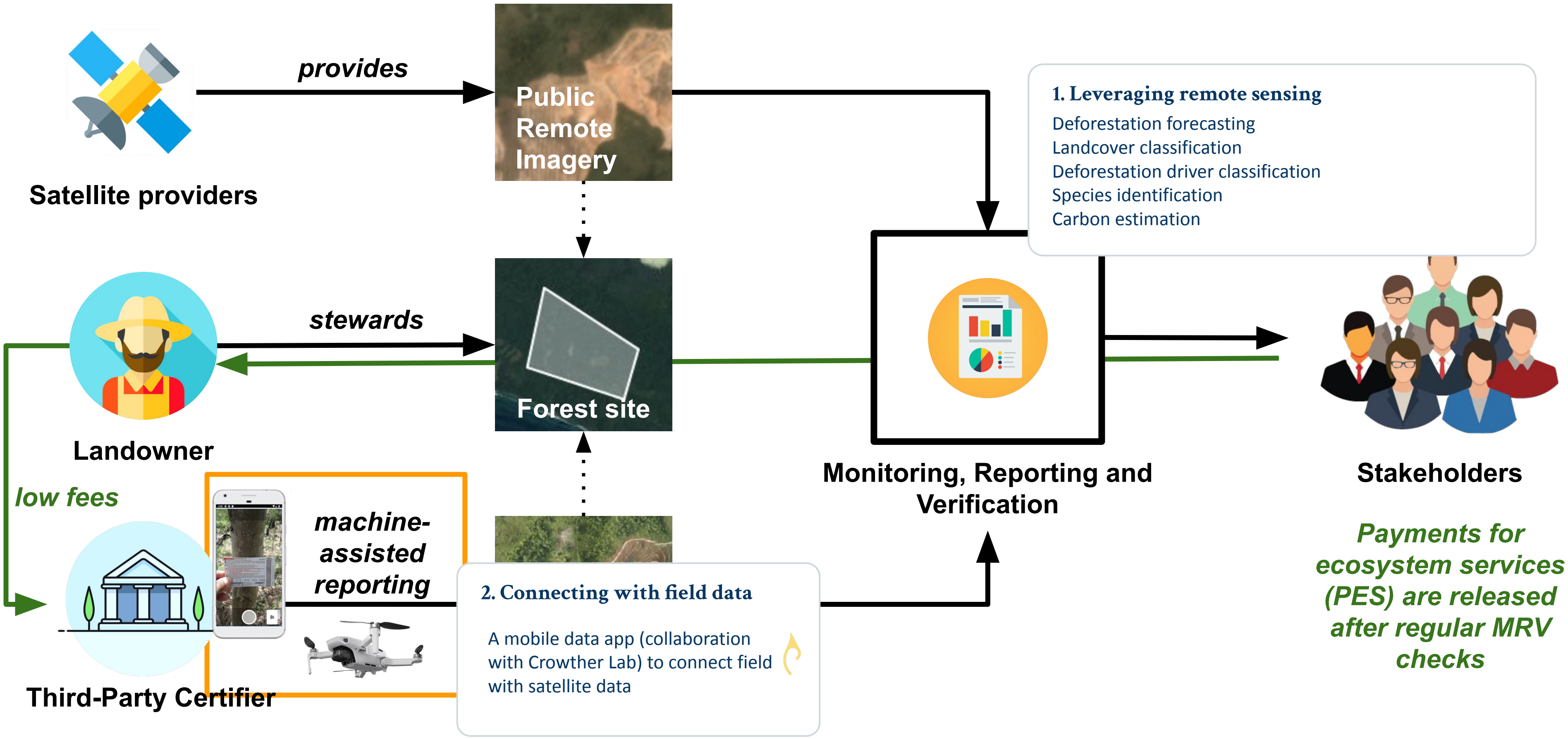
stewards
Forest site
Landowner
expensive fees
manually reports
Field data
Third-Party Certifier
Monitoring, Reporting and Verification
Stakeholders
Payments for ecosystem services (PES) are released after regular MRV checks

Overview



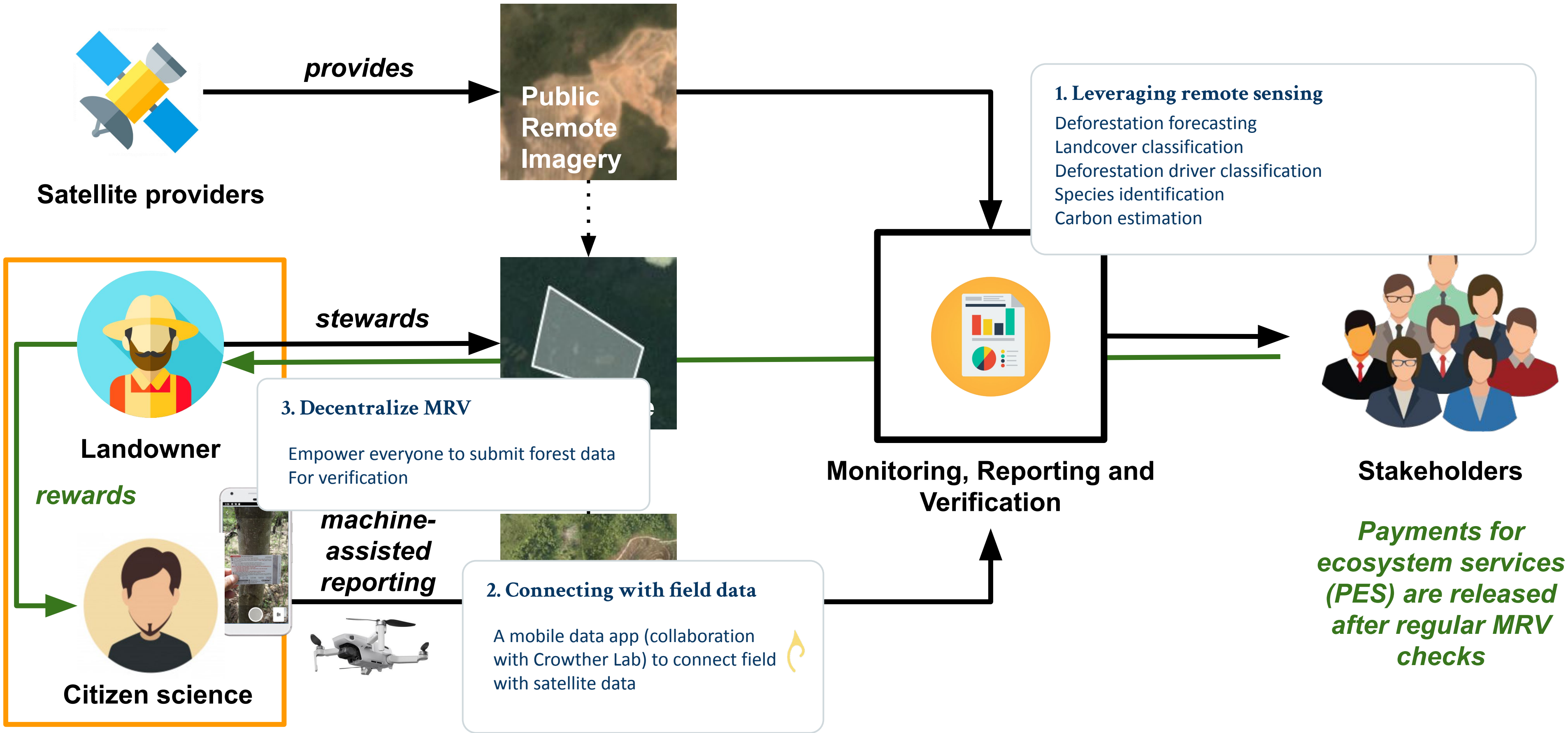
Satellite providers
provides
Public Remote Imagery
1. Leveraging remote sensing
Deforestation forecasting
Landcover classification
Deforestation driver classification
Species identification
Carbon estimation
stewards
Forest site
Landowner
Monitoring, Reporting and Verification
Stakeholders
expensive fees
manually reports
Field data
Third-Party Certifier
Payments for ecosystem services (PES) are released after regular MRV checks

Overview



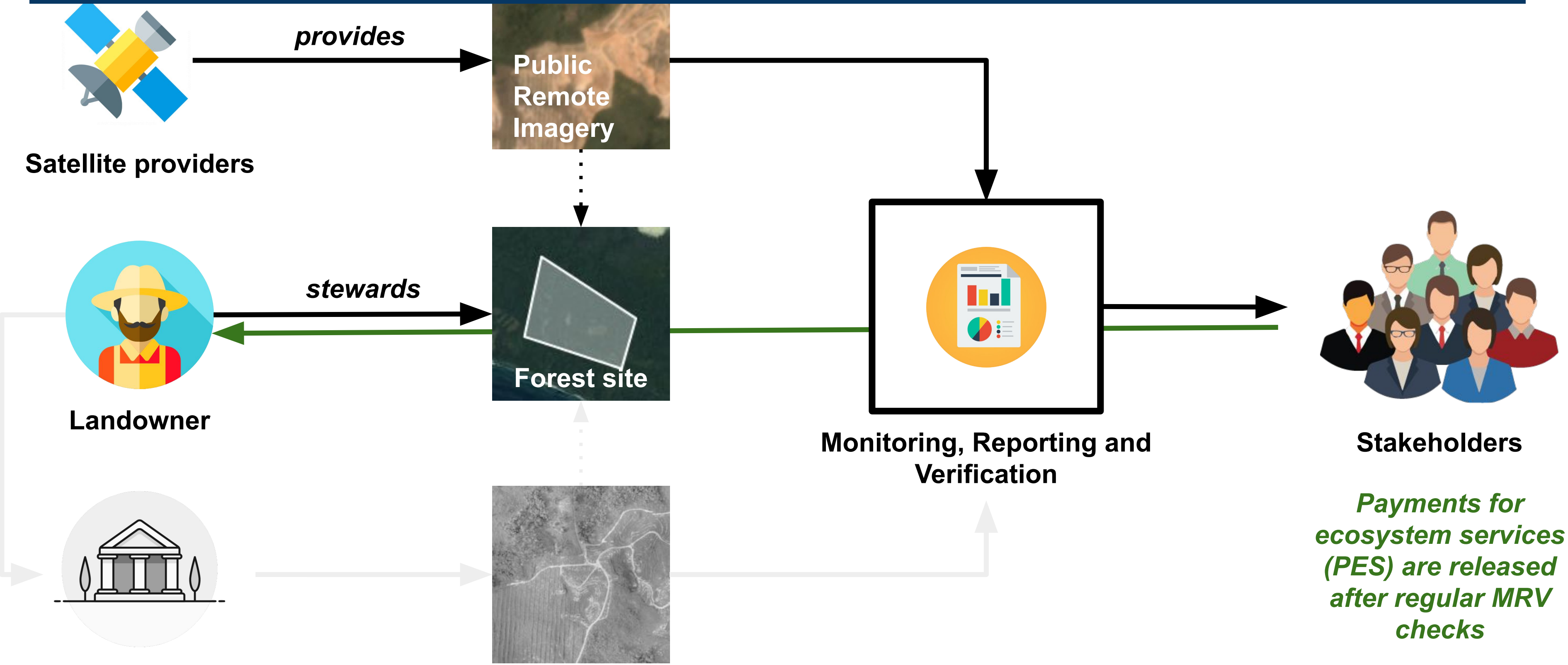
Satellite providers
provides
Public Remote Imagery
1. Leveraging remote sensing
Deforestation forecasting
Landcover classification
Deforestation driver classification
Species identification
Carbon estimation
stewards
Forest site
Landowner
low fees
machine-assisted reporting
Third-Party Certifier
2. Connecting with field data
A mobile data app (collaboration with Crowther Lab) to connect field with satellite data
Monitoring, Reporting and Verification
Stakeholders
Payments for ecosystem services (PES) are released after regular MRV checks

Overview



provides
Public Remote Imagery
Satellite providers
1. Leveraging remote sensing
Deforestation forecasting
Landcover classification
Deforestation driver classification
Species identification
Carbon estimation
stewards
Landowner
rewards
3. Decentralize MRV
Empower everyone to submit forest data
For verification
machine-assisted reporting
Citizen science
2. Connecting with field data
A mobile data app (collaboration with Crowther Lab) to connect field with satellite data
Monitoring, Reporting and Verification
Stakeholders
Payments for ecosystem services (PES) are released after regular MRV checks

Can we leverage remote sensing for MRV?



Utilize multi-modal imagery sources

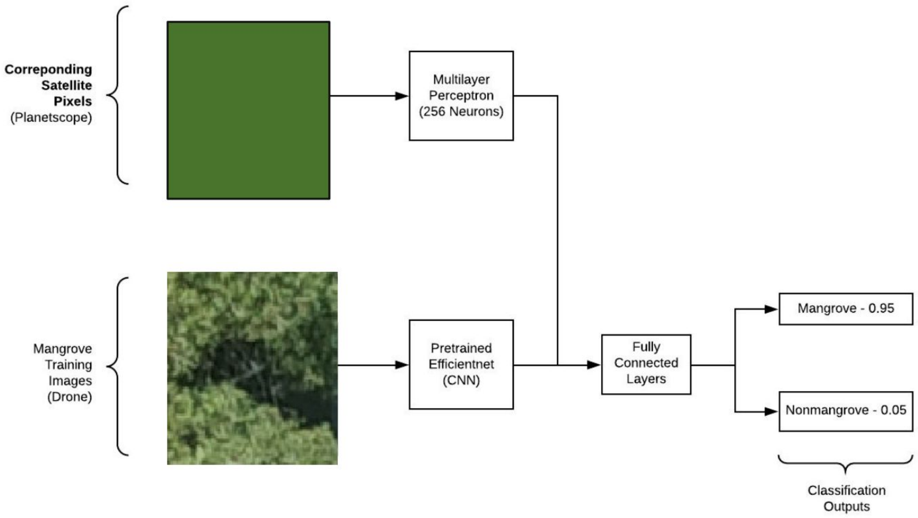
Left: Public data
(Sentinel-2, LANDSAT-8)
>10m/px

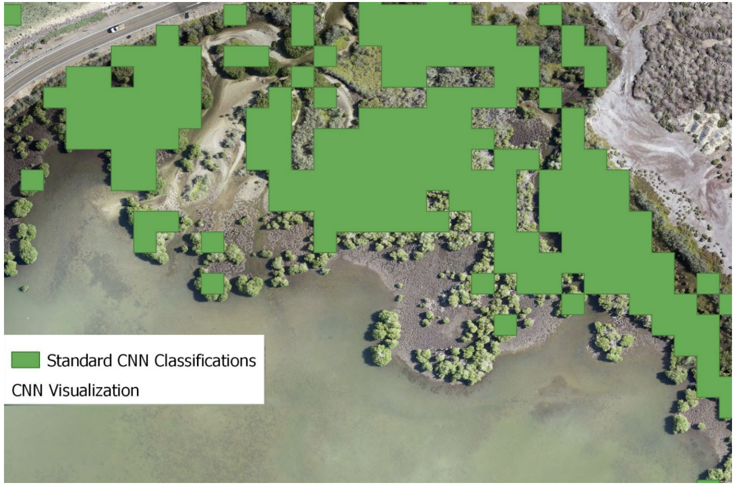
Right: Commercial data
(Planet Labs, MAXAR)
>50cm/px

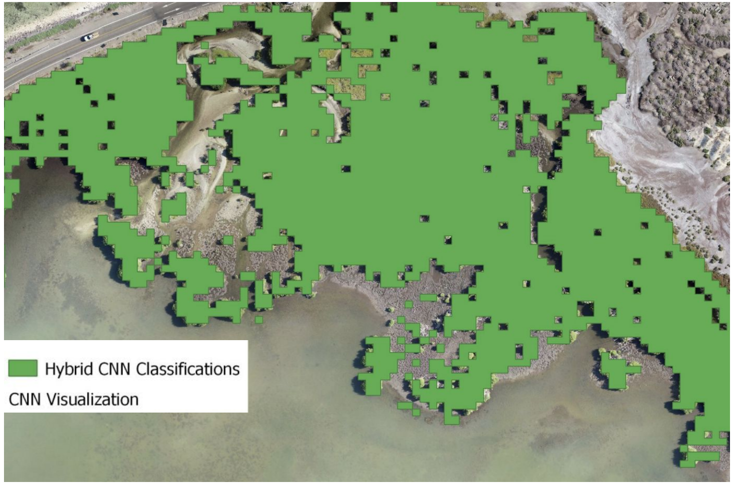
Challenges:
Different sensors, reference systems, access levels and temporal info

Estância Shalom, Brazil

Mangrove classification



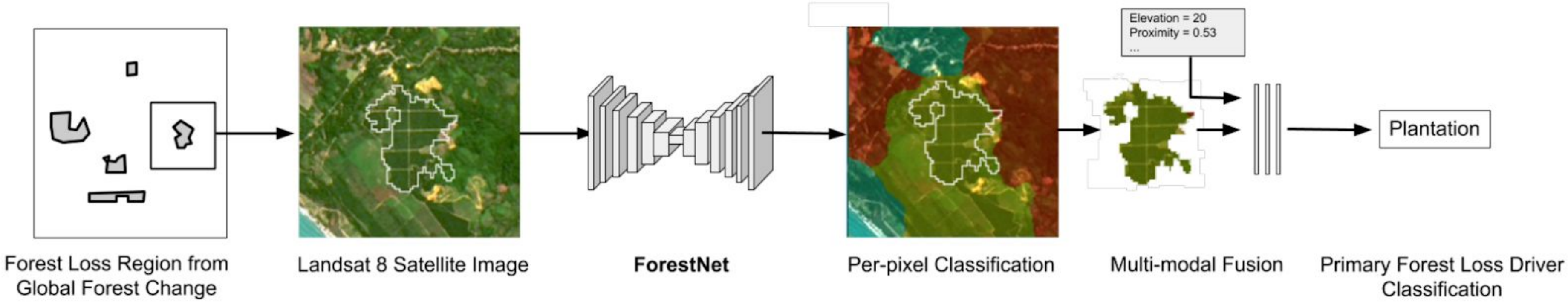




Fusing different data resolutions improves classification accuracy
Challenges: Labeled Data

Mangrove Ecosystem Detection using Mixed-Resolution Imagery with a Hybrid-Convolutional Neural Network
Hicks *et al.*, CCAI ICML20

Deforestation driver classification

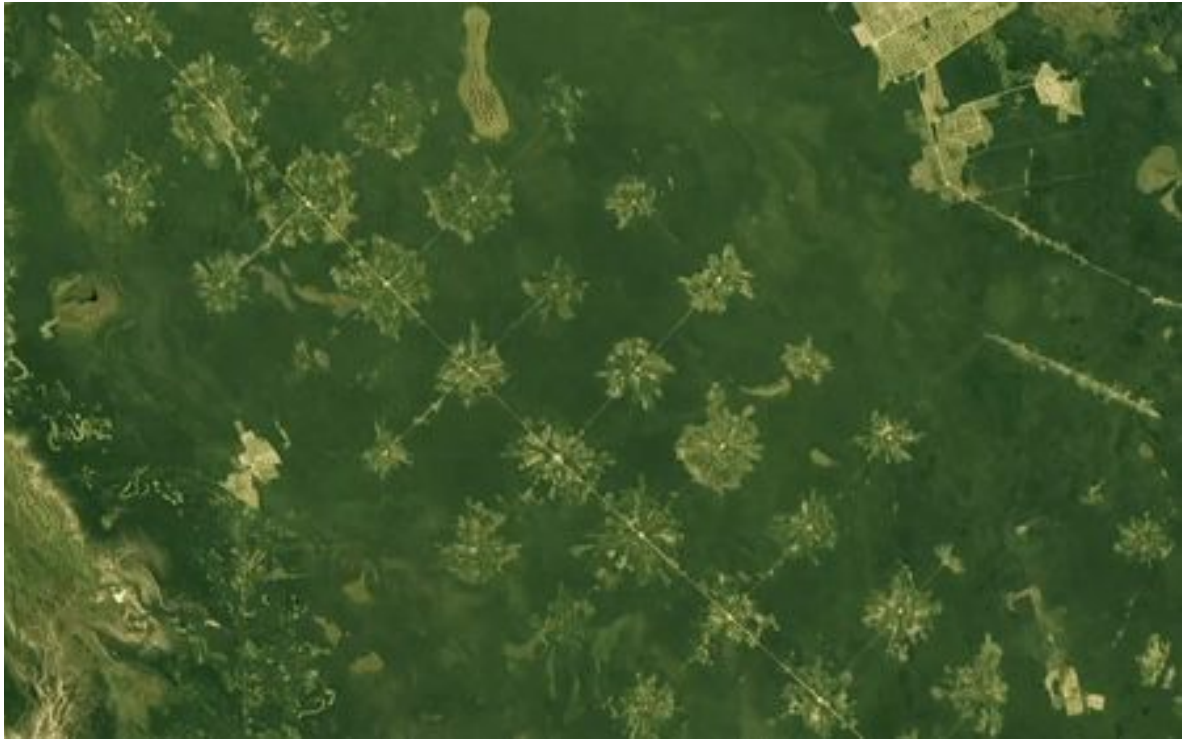


Model	Predictors	Val		Test	
		Acc	F1	Acc	F1
RF	Visible	0.56	0.49	0.49	0.44
RF	Visible + Aux	0.72	0.67	0.67	0.62
CNN	Visible	0.80	0.75	0.78	0.70
CNN + SDA	Visible	0.82	0.79	0.78	0.73
CNN + SDA + PT	Visible	0.83	0.80	0.80	0.74
CNN + SDA + PT	Visible + Aux	**0.84**	**0.81**	**0.80**	**0.75**

Fusing multi-modal data sources improve classification accuracy
Challenges: Labeled Data

ForestNet: Classifying Drivers of Deforestation in Indonesia using Deep Learning on Satellite Imagery
Irvin *, Sheng * *et al.* CCAI NeurIPS20

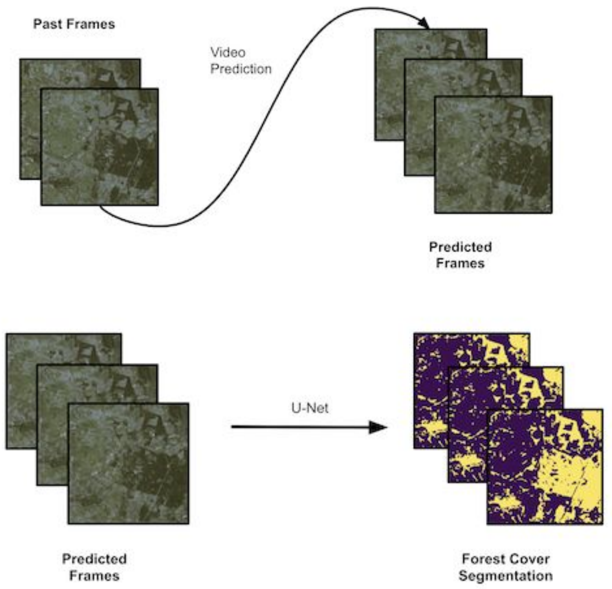
Utilize large time series



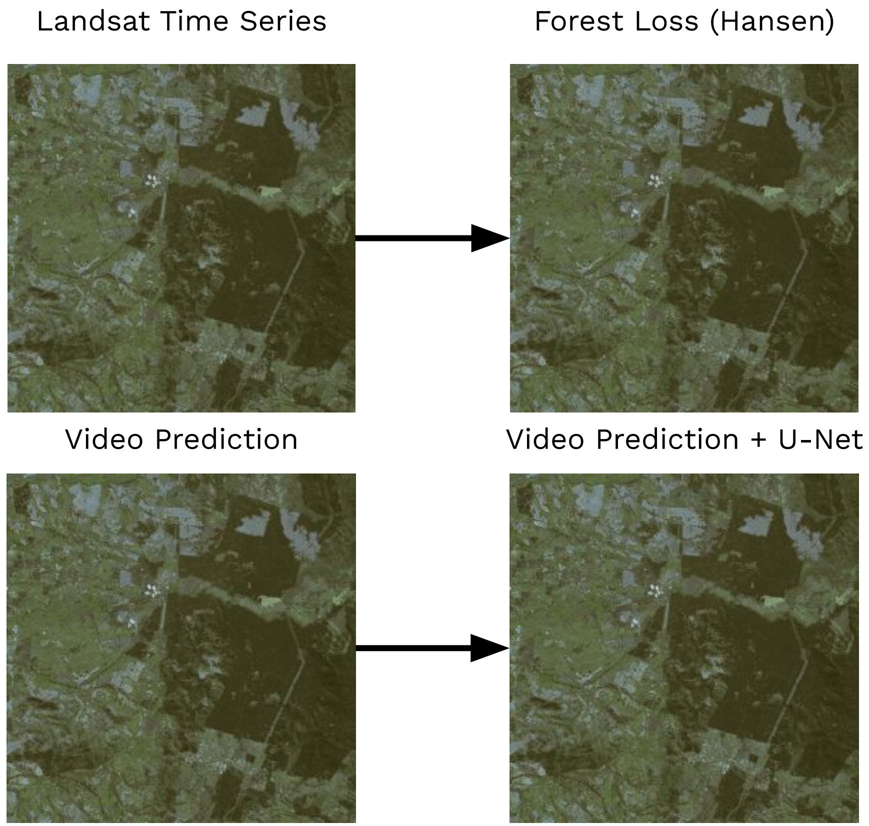
Challenges:
Largely unlabeled data:
E.g. deforestation exhibits visually recognizable patterns

Irregular temporal steps:
Visual data depends on satellite revisiting rate and cloudfree image

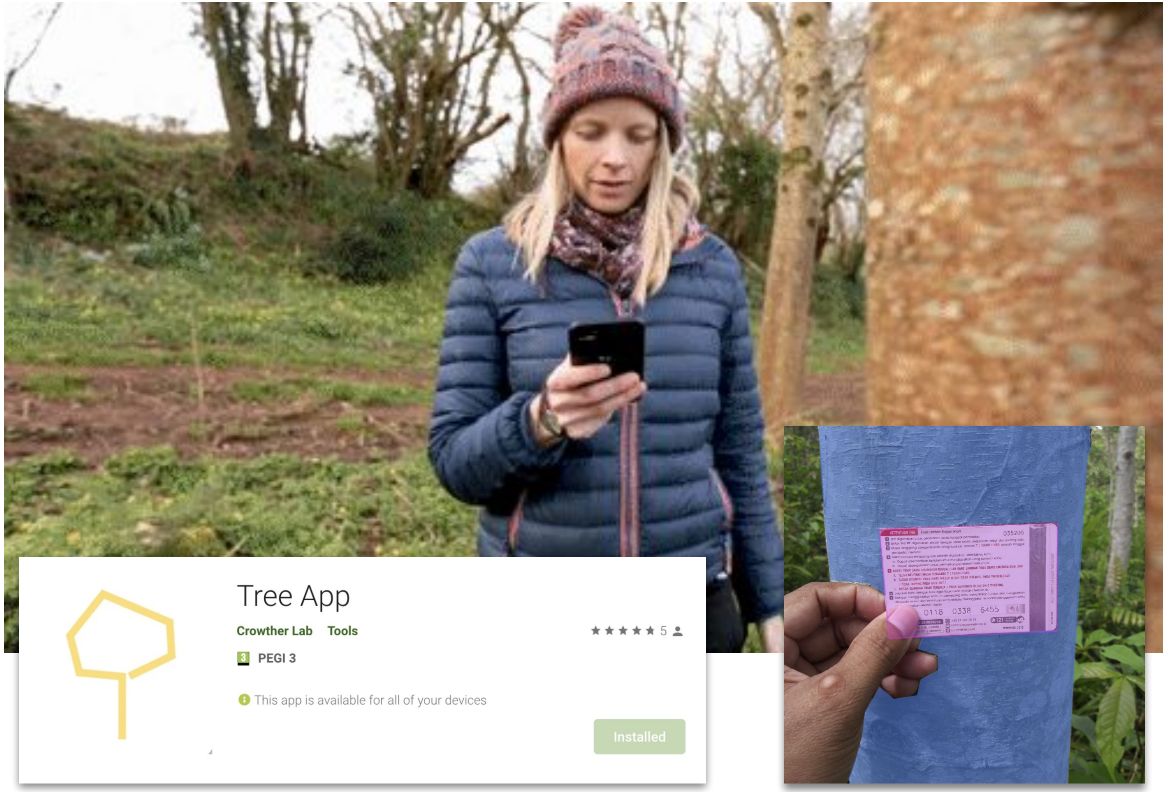
Deforestation forecasting with self-supervision



Challenges:
Difficult to train
Needs lots of
computational resources



Machine-Learning Based Allometric Estimation

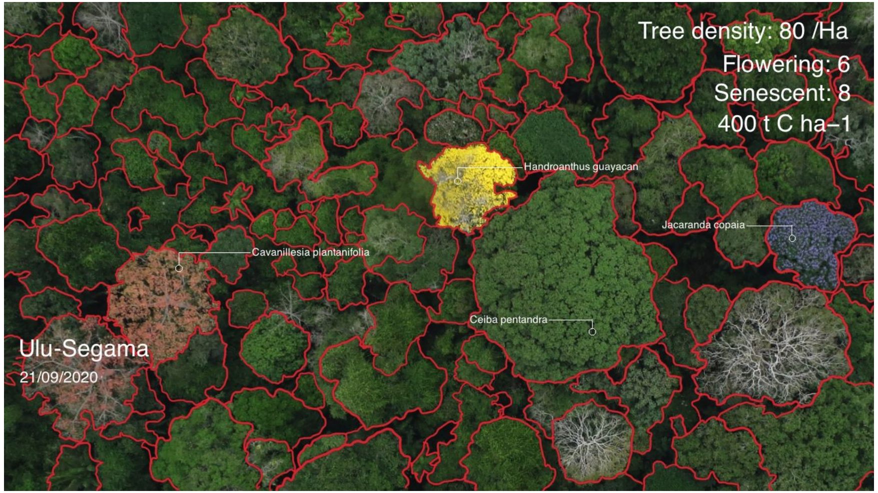


TreeApp: Predicting forest properties through a smartphone camera and citizen science. S. Max et al. (in preparation)

Tree App:
Estimating diameter at breast height (DBH) and species automatically from a single image

Challenges:
Noisy GPS
Low-cost and offline models
Collecting data still tedious

Drone-Based Biomass Estimation



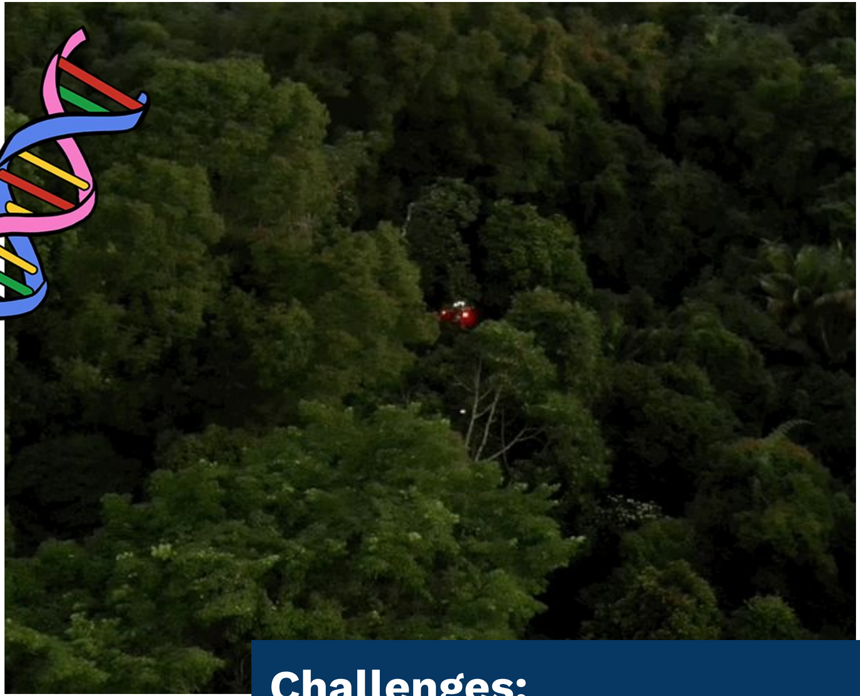
Challenges:
Limited labeled data for model training

Drone-Based Biodiversity Measurements

Work from our collaborator Prof. Stefano Mintchev in "Environmental Robotics"

Challenges:
Difficult Terrain
Network Connection

Drone-Based Environmental DNA Sampling



Sample Strategies:
eDNA through water filtration
eDNA through surface collection
eDNA through air filtration

Challenges:
Collecting enough eDNA
Dense Canopy

Joint work displayed during the XPRIZE semifinal in Singapore (Deiner, Mintchev, Dao *et al.*)

Limits of computation-centric approaches & Role of local communities

Systematic Overestimation Study

- AGB dataset from agro-forestry sites
- 4663 trees, 28 species and 3.17 ha
- Each tree registered with DBH, species, and GPS location
  RGB drone images per site

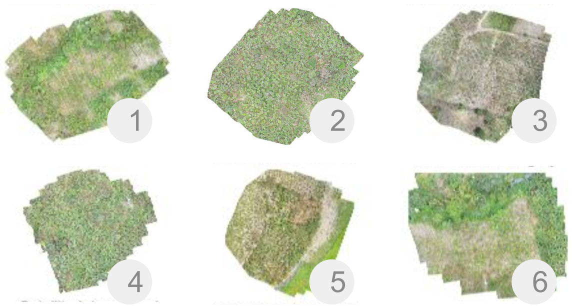


Fig. 1 Information about each site

SITE NO.	NO. OF TREES	NO. OF SPECIES	PLOT AREA	AGB DENSITY
1	743	17	0.53	19
2	929	19	0.47	32
3	789	21	0.51	26
4	484	13	0.56	16
5	872	15	0.62	24
6	846	16	0.48	27

Equations (1) and (2): Allometric equations from [8] and [9]

$$log_{10}AGB_{standard} = -0.834 + 2.223(log_{10}DBH)$$

$$AGB_{musacea} = 0.030 * DBH^{2.13}$$

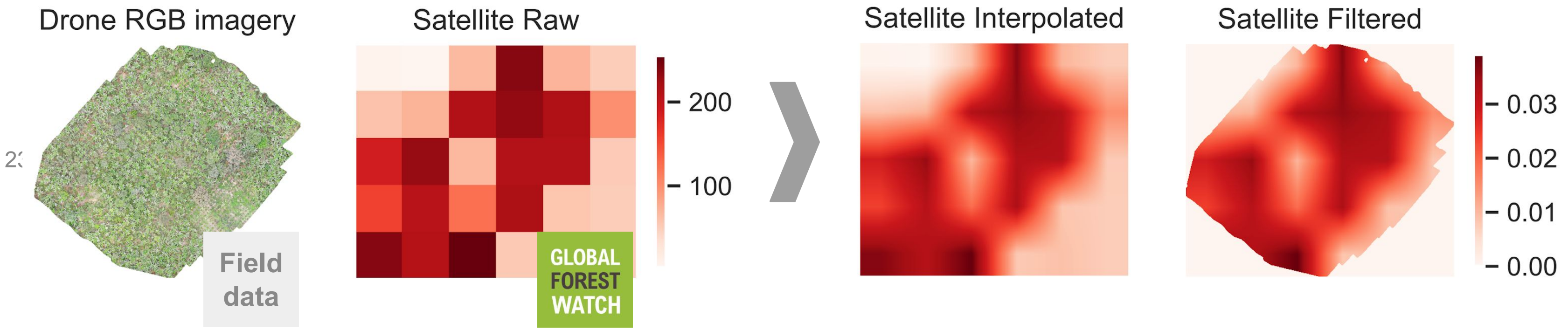
"Tackling the Overestimation of Forest Carbon with Deep Learning and Aerial Imagery"
Reiersen *et al.*, CCAI ICML 2021

Benchmarking satellite-based AGB density estimation against field data

Global Forest Watch product: *Aboveground live woody biomass density* [10]

- 30mx30m resolution, 70k GLAS observations with deep learning model
- Lidar-derived canopy metrics and region-specific allometric equations

Map interpolated and filtered on the locations for all field data sites:



"Tackling the Overestimation of Forest Carbon with Deep Learning and Aerial Imagery"
Reiersen *et al.*, CCAI ICML 2021

The satellite-based estimates significantly overestimates AGB density by a factor of 10

- The AGB density (kg/ha) per polygon was overestimated for **all of the 6 sites** with a factor ranging up to 10 times the field data

Site No.	Ground Truth	Filtered	Over Estimation
1	19	176	$\times 9.2$
2	27	160	$\times 5.9$
3	24	47	$\times 2.0$
4	24	62	$\times 2.6$
5	17	19	$\times 1.1$
6	29	141	$\times 4.9$

Fig. 2 AGB density (kg/ha) of the field data (Ground truth) and of the satellite based estimations (Filtered)

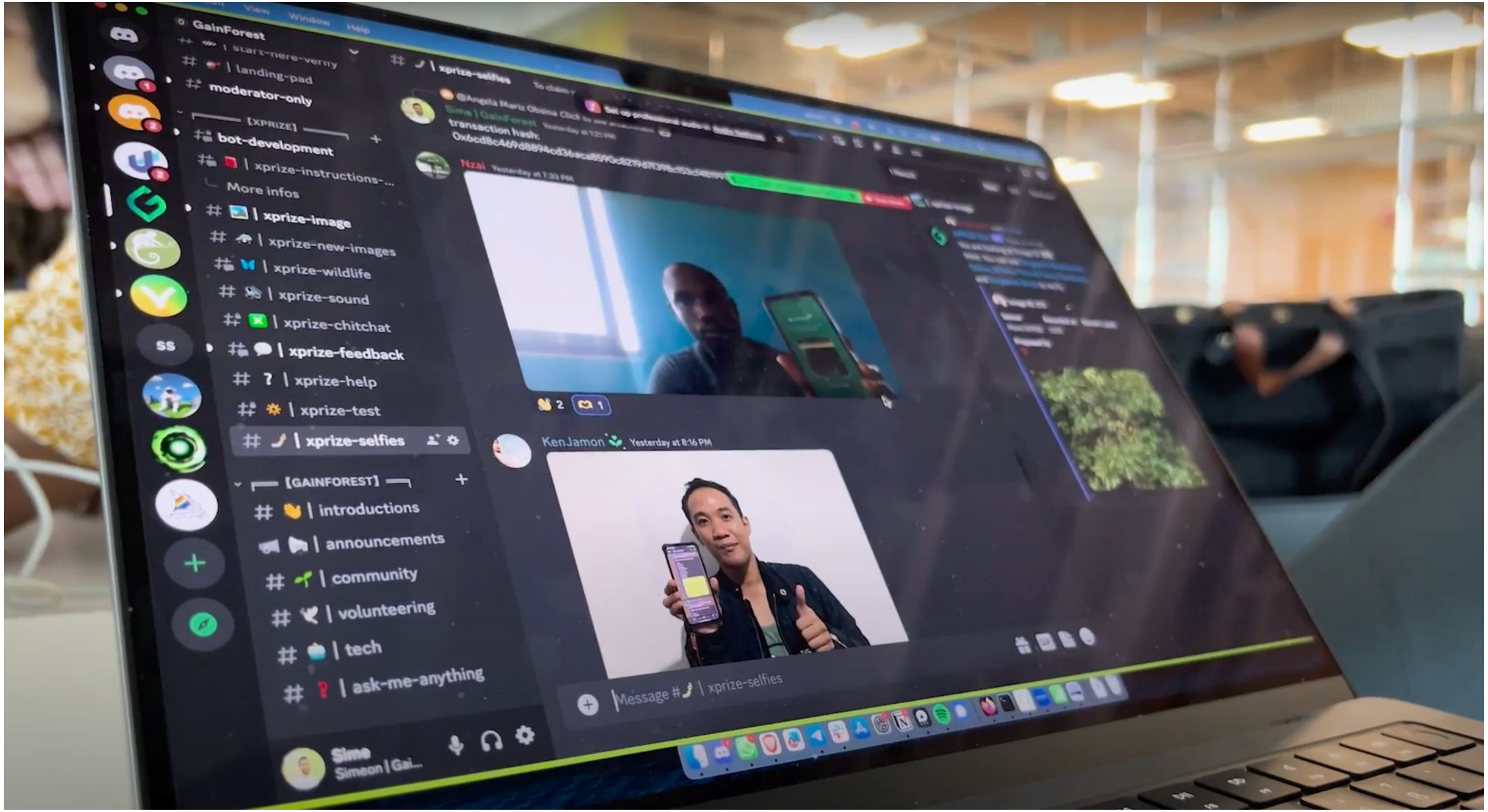
"Tackling the Overestimation of Forest Carbon with Deep Learning and Aerial Imagery"
Reiersen *et al.*, CCAI ICML 2021

The crucial role of Indigenous communities in MRV

Participatory
Mapping

Strategy
Games

Citizen Science for Conservation



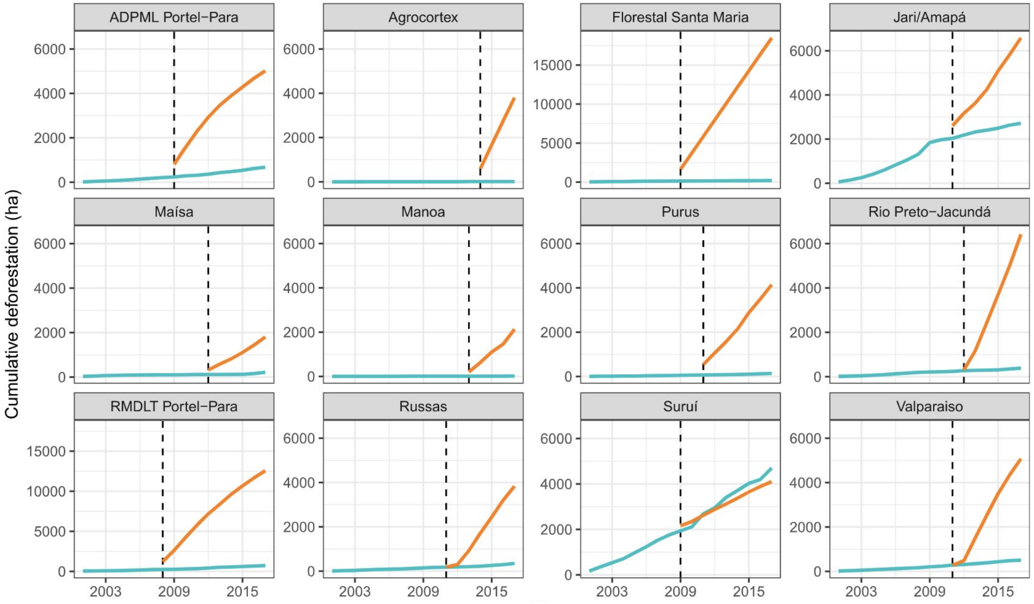
Indigenous modeling in REDD+ in Brazil (Quiz ❓)

Label

Baselines provided by project

Deforestation in synthetic controls

Only one out of 12 forest carbon projects used a model that was co-designed with Indigenous communities. **Can you guess which one?**



"Overstated carbon emission reductions from voluntary REDD+ projects in the Brazilian Amazon"
West *et al.*, PNAS, 2021

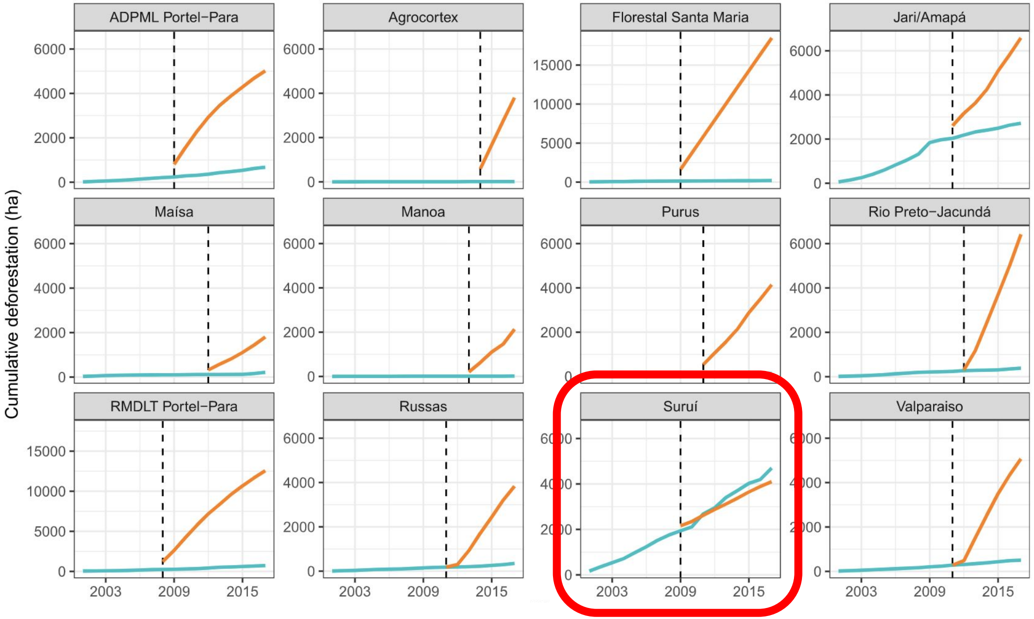
Indigenous modeling in REDD+ in Brazil

Label

Baselines provided by project

Deforestation in synthetic controls

Only one out of 12 forest carbon projects used a model that was co-designed with Indigenous communities. **Can you guess which one?**



"Overstated carbon emission reductions from voluntary REDD+ projects in the Brazilian Amazon"
West *et al.*, PNAS, 2021


Number of citizen scientists that identified at least one species: 30

Lesson Learned

Local communities

must be involved in the co-design of impactful technology

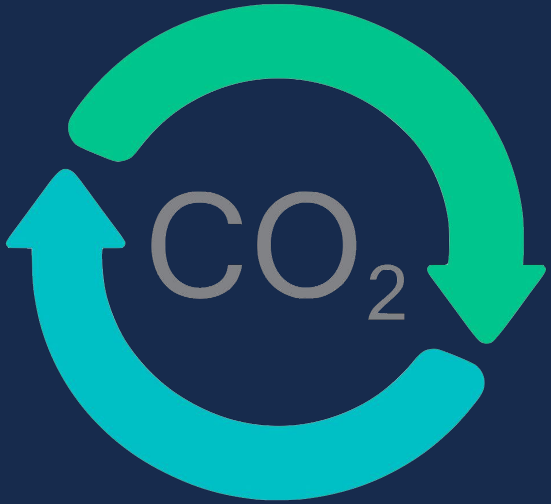
need to be one of the main beneficiaries of impactful technology

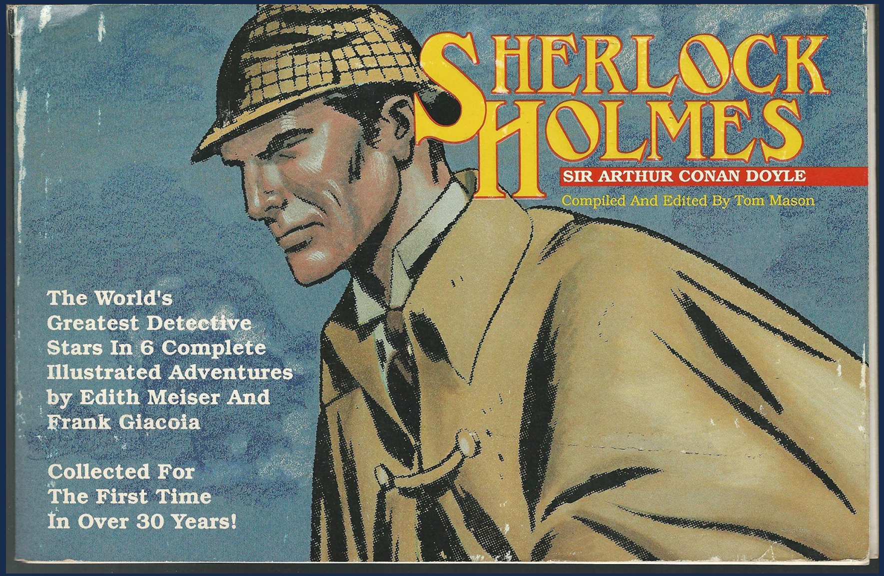
Practice

Carbon Project Detective

GainForest: Making ecological data transparent

🕵️‍♀️ *Forest Carbon Detective Hands-On*

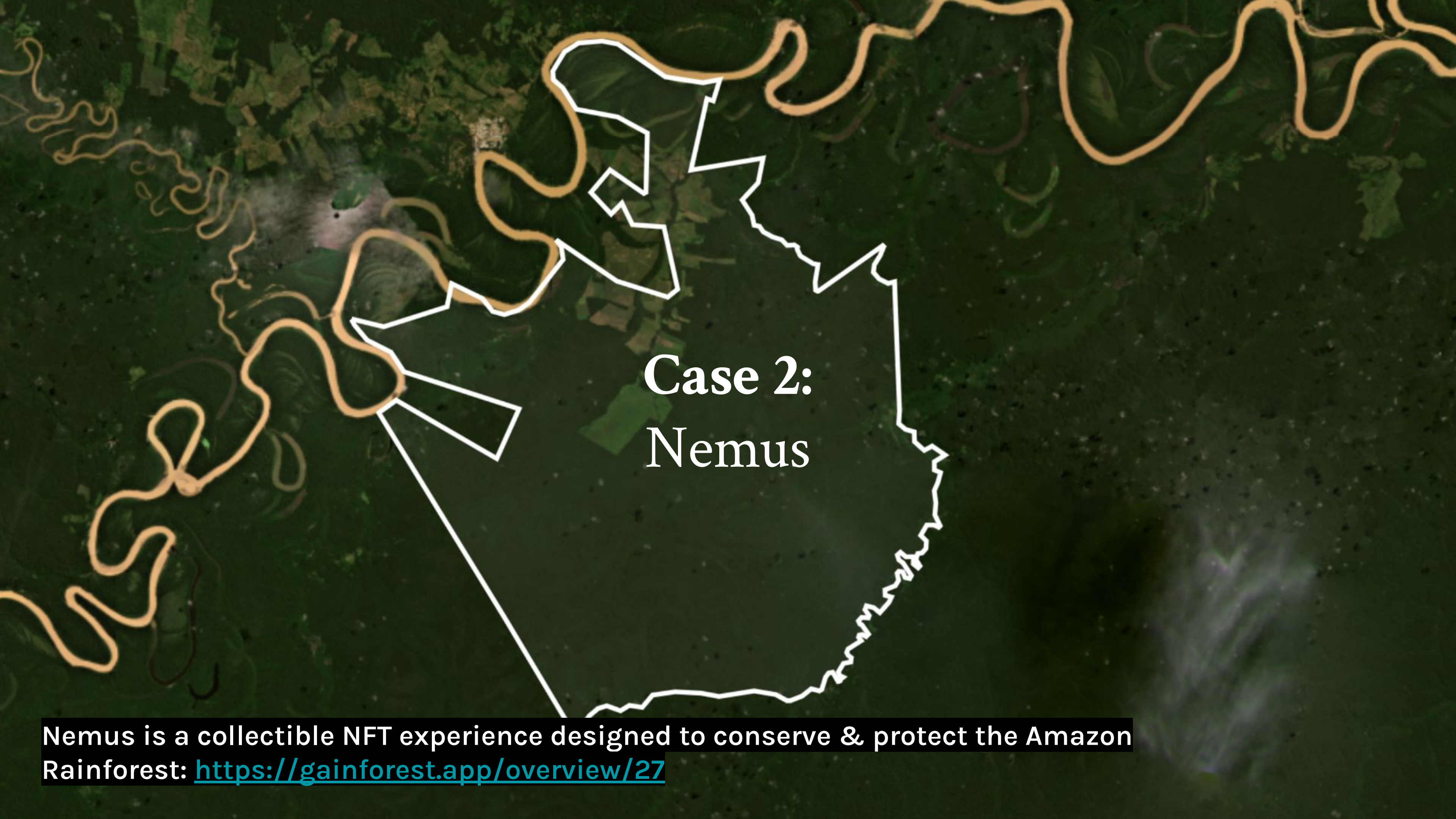




Case 1:

Weyerhaeuser Reforestation Project

Verra Project 960, “WEYERHAEUSER URUGUAY” FOREST PLANTATIONS ON DEGRADED GRASSLANDS UNDER EXTENSIVE GRAZING: https://gainforest.app/overview/8



Case 2: Nemus

Nemus is a collectible NFT experience designed to conserve & protect the Amazon Rainforest: https://gainforest.app/overview/27

Case 3:
Kariba REDD+ Project

VCS-902, The Kariba REDD+ Project will generate approximately 196,500,000 carbon credits from reduced emissions associated with deforestation over 30 years:
https://gainforest.app/overview/32

Lessons learned

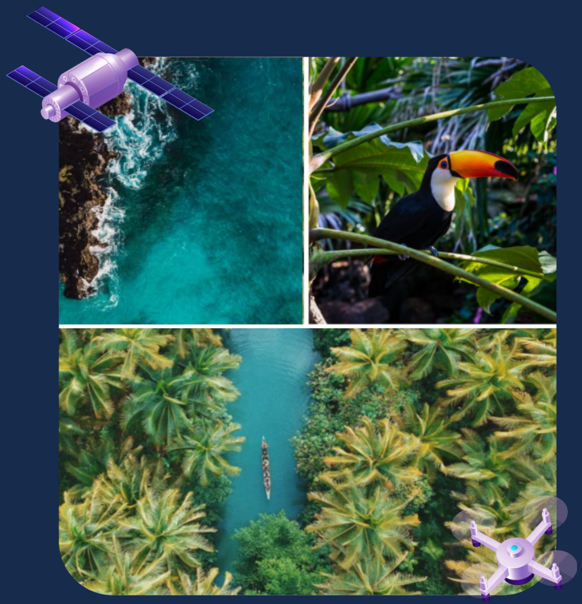
- **Be <u>transparent and open</u>**
  - MRV algorithms and data influence policy and who receives payments or not, and can create dangerous feedback loops
- **Co-design with <u>local communities</u>**
  - Data ownership
  - Scenario-specific metrics
  - Otherwise we build tools that are satisfying us and nobody else
- **Research <u>interdisciplinary</u>**
  - MRV is about human behavior
  - Modeling real world requires ML to expand interdisciplinary
  - Goodhardt's Law: "When a measure becomes a target, it ceases to be a good measure"

Thank you!

David Dao

@dwddao

@gainforestnow



Appendix

More research, etc …

Challenge

Automated forest validation opens up possibility of *untruthfully* reported imagery

Attack vectors

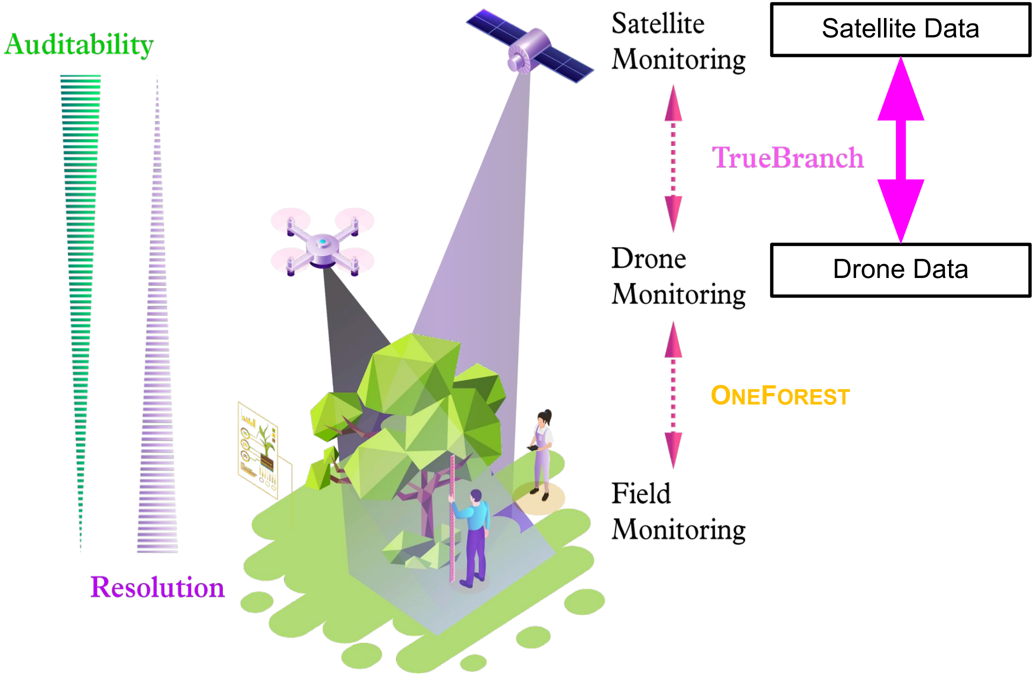
Reported Land-Use

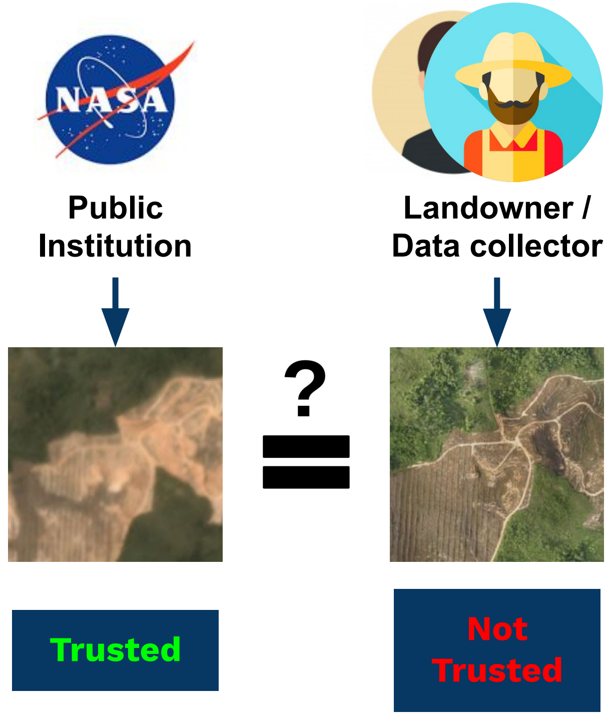



	true time **true** location	**wrong** time **true** location	**true** time **wrong** location	**modified image**
Detected Forest Cover	**medium**	**high**	**high**	**high**

Simona Santamaria*, David Dao*, Björn Lütjens*, TrueBranch: Metric Learning-Based Verification of Forest Conservation Projects, 2020 ICML Workshop (CCML), *equal contribution, https://arxiv.org/abs/2004.09725

Novel opportunities through data fusion

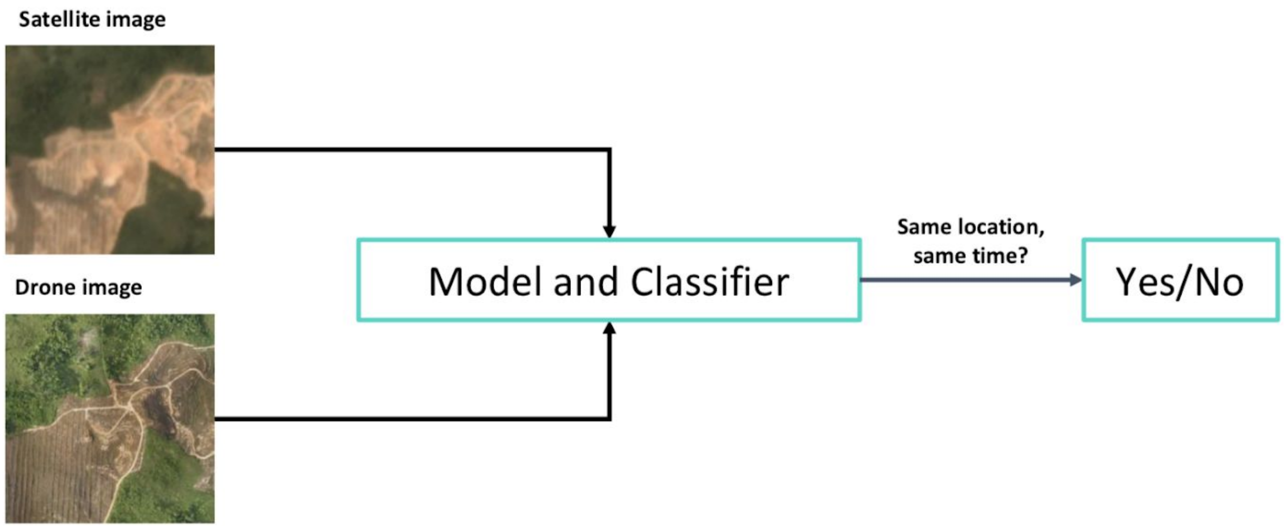




Classifying truthfulness

How to distinguish truthful imagery from untruthful imagery?

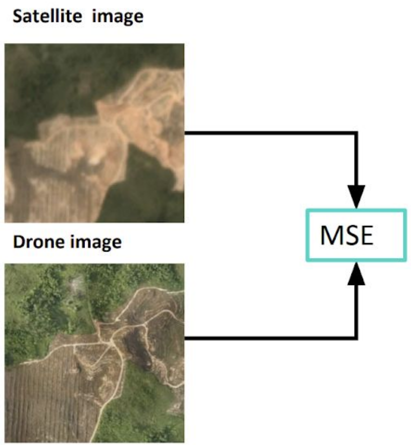
- Image Registration: Matching Drone images with Satellite images



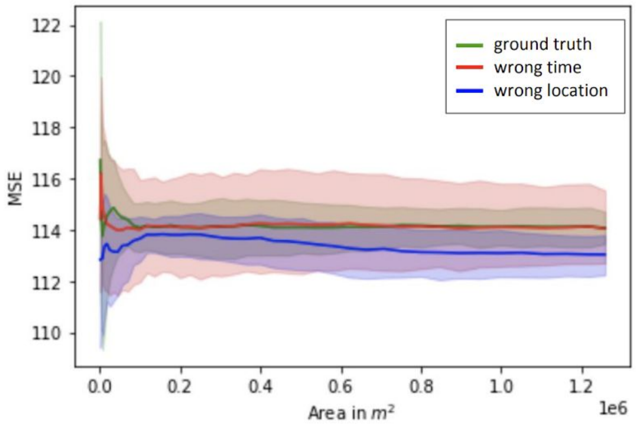
Simona Santamaria*, David Dao*, Björn Lütjens*, TrueBranch: Metric Learning-Based Verification of Forest Conservation Projects, 2020 ICML Workshop (CCML), *equal contribution, https://arxiv.org/abs/2004.09725

Classifying truthfulness

- Nominal distance metrics of MSE in pixels space



$$MSE = \frac{1}{mn} \sum_{i=0}^{n-1} \sum_{j=0}^{m-1} [A(i,j) - B(i,j)]^2$$

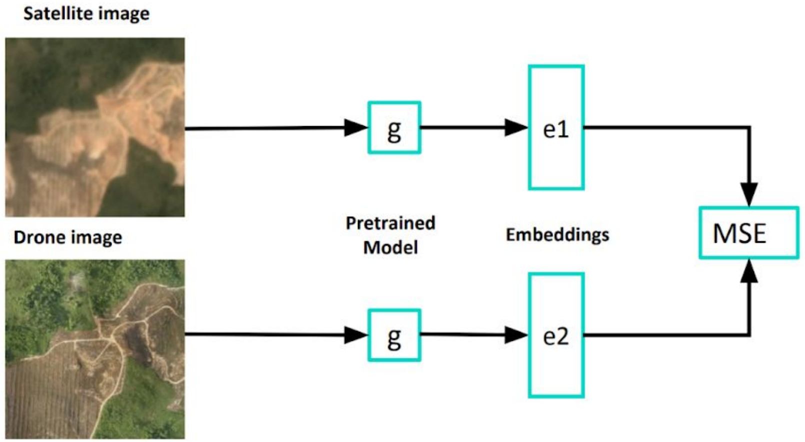


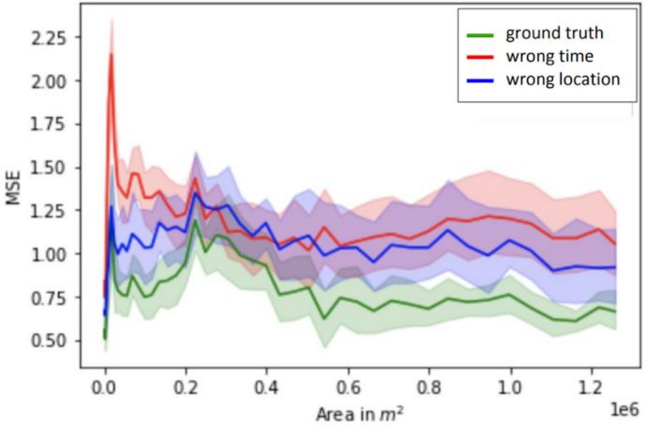
Simona Santamaria*, David Dao*, Björn Lütjens*, TrueBranch: Metric Learning-Based Verification of Forest Conservation Projects, 2020 ICML Workshop (CCML), *equal contribution, https://arxiv.org/abs/2004.09725

Classifying truthfulness

- Nominal distance metrics of MSE in feature space

$$MSE = \frac{1}{mn}\sum_{i=0}^{n-1}\sum_{j=0}^{m-1}[A(i,j) - B(i,j)]^2$$

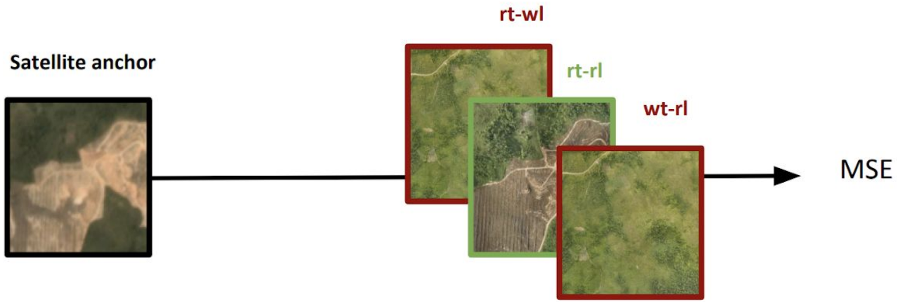


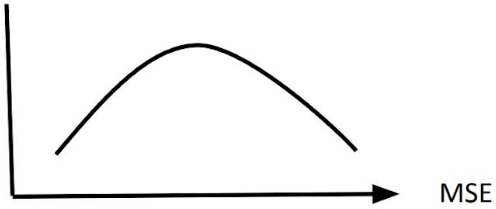


Simona Santamaria*, David Dao*, Björn Lütjens*, TrueBranch: Metric Learning-Based Verification of Forest Conservation Projects, 2020 ICML Workshop (CCML), *equal contribution, https://arxiv.org/abs/2004.09725

Classifying truthfulness

- MSE in pixel space and RESISC-45 feature space not sufficient

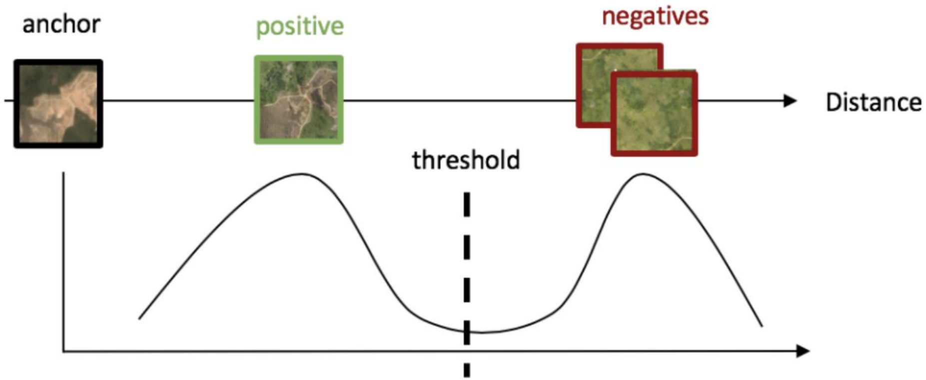




Simona Santamaria*, David Dao*, Björn Lütjens*, TrueBranch: Metric Learning-Based Verification of Forest Conservation Projects, 2020 ICML Workshop (CCML), *equal contribution, https://arxiv.org/abs/2004.09725

Metric Learning

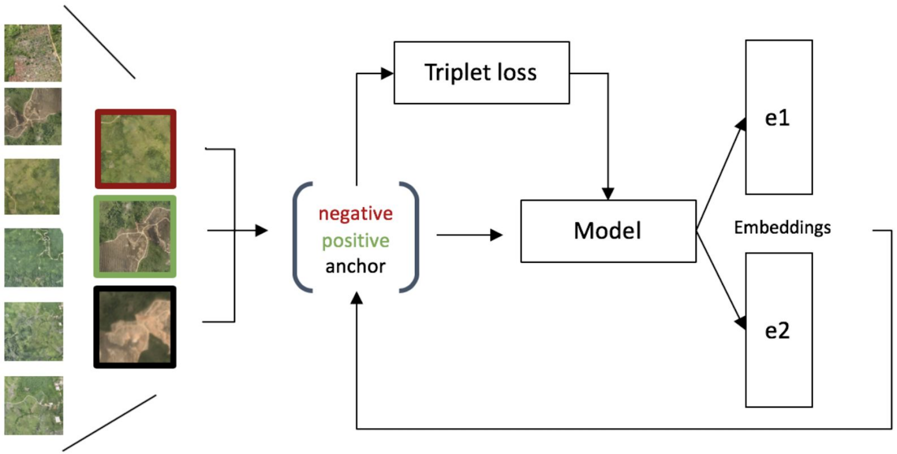
The distance between the anchor and positive image is decreased while the distance between the anchor and negative image is increased.

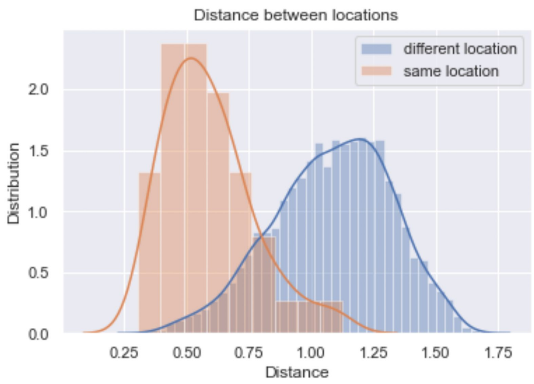


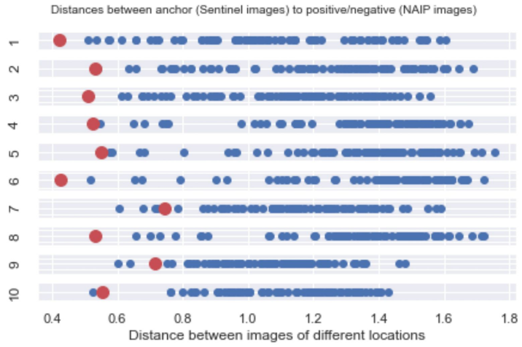
$$L(a, p, n) = max(|f(a) - f(p)|^2 - |f(a) - f(n)|^2 + \alpha, 0)$$

Simona Santamaria*, David Dao*, Björn Lütjens*, TrueBranch: Metric Learning-Based Verification of Forest Conservation Projects, 2020 ICML Workshop (CCML), *equal contribution, https://arxiv.org/abs/2004.09725

TrueBranch: Metric Learning-based Verification







TrueBranch enables the verification of truthfully reported drone imagery from untrusted parties

Simona Santamaria*, David Dao*, Björn Lütjens*, TrueBranch: Metric Learning-Based Verification of Forest Conservation Projects, 2020 ICML Workshop (CCML), *equal contribution, https://arxiv.org/abs/2004.09725