# Data ethics for monitoring reporting and verification service: A study of deforestation detection of Coffee supply chains in South America

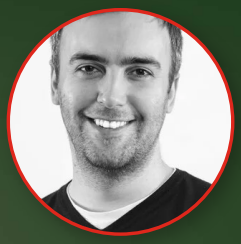
Written by Rob Maslin, Satellite Applications Catapult,
November 2022 (updated for public release October 2023)

We work with

UKRI Innovate UK

CATAPULT Satellite Applications

## WRITTEN BY



**Rob Maslin,**

**Design Lead**
Satellite Applications Catapult

View Rob Maslin
Linkedin® Profile

With contributions from Justin Byrne on Data Analysis

The report is written for completion of the Open Data Institute's Data Ethics certification.

The report is based on the work of the ForestMind project team at Satellite Applications Catapult:

Simon Kartar
Rob Maslin
Nasreen Dhanji
Sarah Cheesbrough
Jade Saunders
Chris McNeil
Chris Hobbs
Ray Clacher
Justin Byrne

# Table of Contents

# Acronyms

| | |
|---|---|
| **AOI** | Area of interest |
| **UoL** | University of Leicester |
| **MRV** | Monitoring Verification and Reporting |
| **IPCC** | International Panel in climate change |

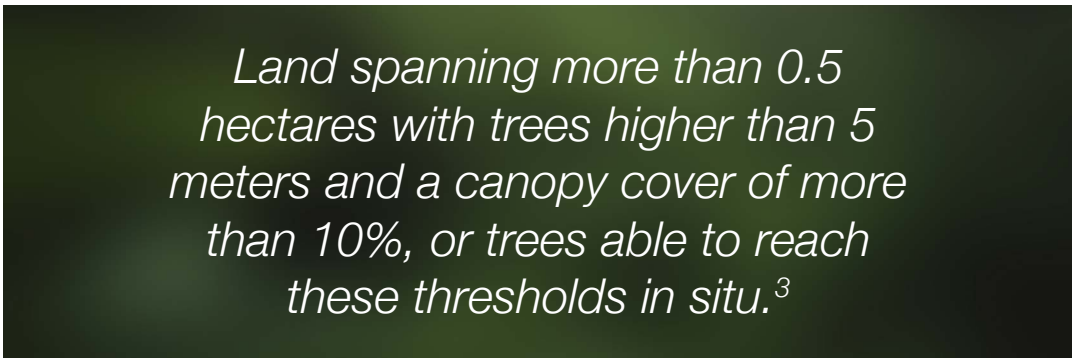
# 1. BACKGROUND

## 1.1 Challenge of Deforestation

ForestMind is a project delivered by Satellite Applications Catapult and funded by ESA to pilot real-world demonstration of remote sensing services to detect deforestation in commodity supply chains. These services are commonly referred to as Monitoring Verification and Reporting (MRV) services. For the purpose of this data ethics study of MRV services, the focus is a coffee brand wanting to understand more about deforestation risks in their supply chain.

The demand for MRV services is driven by the recognition from the International Panel in Climate Change that deforestation[1] causes 23% of anthropogenic greenhouse gases. Deforestation in the coffee supply chain could potentially risk net-zero commitments, cause reputational brand damage, and limit any organisation in the supply chain's access to export to international markets regulated under emerging environmental and deforestation legislation. Any organisation trading coffee in the UK will need ensure it can meet the incoming UK regulations[2], which require undertaking due diligence on supply chains deforestation footprint and stopping commodities responsible for deforestation from entering the UK. The cut-off date for historical deforestation (accepted as a baseline) is likely to be the end of December 2020.

Legislation is likely to follow the definition of deforestation described by the Food and Agriculture Organisation (FAO) at the UN:




By delivering MRV services with real world insights, the ForestMind project has been able to observe and assess the practicalities, limitations and ethical responsibilities. The project does not intend to deliver any further commercial services after the pilot demonstration. This paper intends to shine a light on a small number of legitimate risks and unintended impacts of MRV services. Future activities building on the project will be to use the learnings to support the emergence of an effective responsible MRV services business ecosystem.

Outside the scope of the study are the important in-depth ethical considerations of worker conditions in supply chains as well as the full environmental and carbon impacts.

[1] *Valérie Masson-Delmotte et al, Climate Change and Land: IPCC, An IPCC Special Report on climate change, desertification, land degradation, sustainable land management, food security, and greenhouse gas fluxes in terrestrial ecosystems – Summary for Policymakers, revised (2020), p8*

[2] *https://www.gov.uk/government/news/government-sets-out-plans-to-clean-up-the-uks-supply-chains-to-help-protect-forests*

[3] *Global Forest Resources Assessment: Terms and Definitions, Food and Agriculture Organisation, United Nations, (2020), p4*

# 2. Who Is Affected?

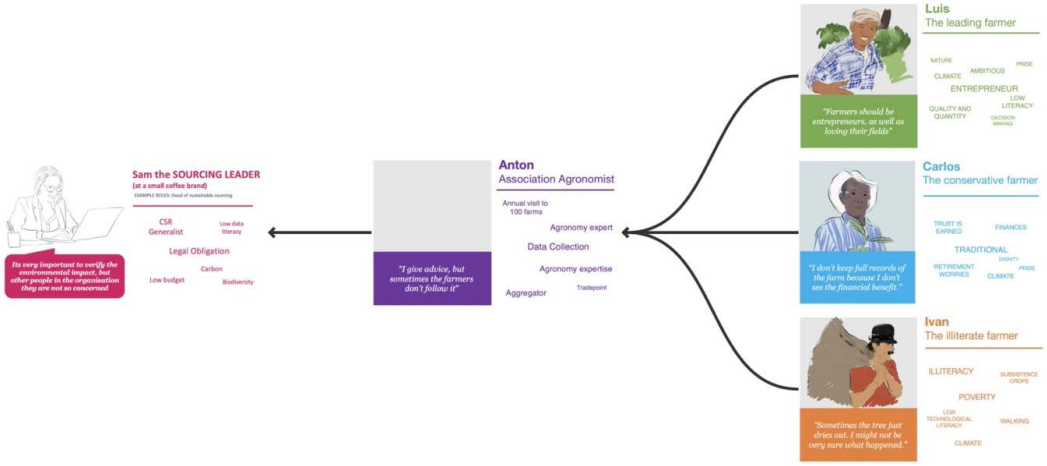



*Figure 1: Persona's developed from research engagements with coffee supply chains*

The complexity and stakeholders in supply chains can vary considerably according to commodities and the business models of the stakeholders involved. For the illustrative purpose of this study, a relatively simple coffee supply chain has been chosen and illustrated above. The data here is taken from real-world studies and the piloted ForestMind service. Research was undertaken in LATAM, where meetings with coffee farmers revealed a range of backgrounds and capabilities from a small number who demonstrate good business acumen to illiterate[4] smallholder and subsistence farmers living in poverty.

Farmers in this supply chain belong to a farmers' association (referred to as associations); they pay an annual fee, and the association provides a limited number of agronomists to visit farms. The association buys coffee from the farmers and negotiates trade agreements with coffee brands like ForestMind's customer.

ForestMind and the farmer associations' primary contact at the coffee brand is the head of sustainable sourcing, who is the main CSR/ESG person in the organisation and the person responsible for visiting the farmers' association and a selection of suppliers annually.

[4] *http://uis.unesco.org/en/country/gt?wbdisable=true*

# 3. The Forestmind Service

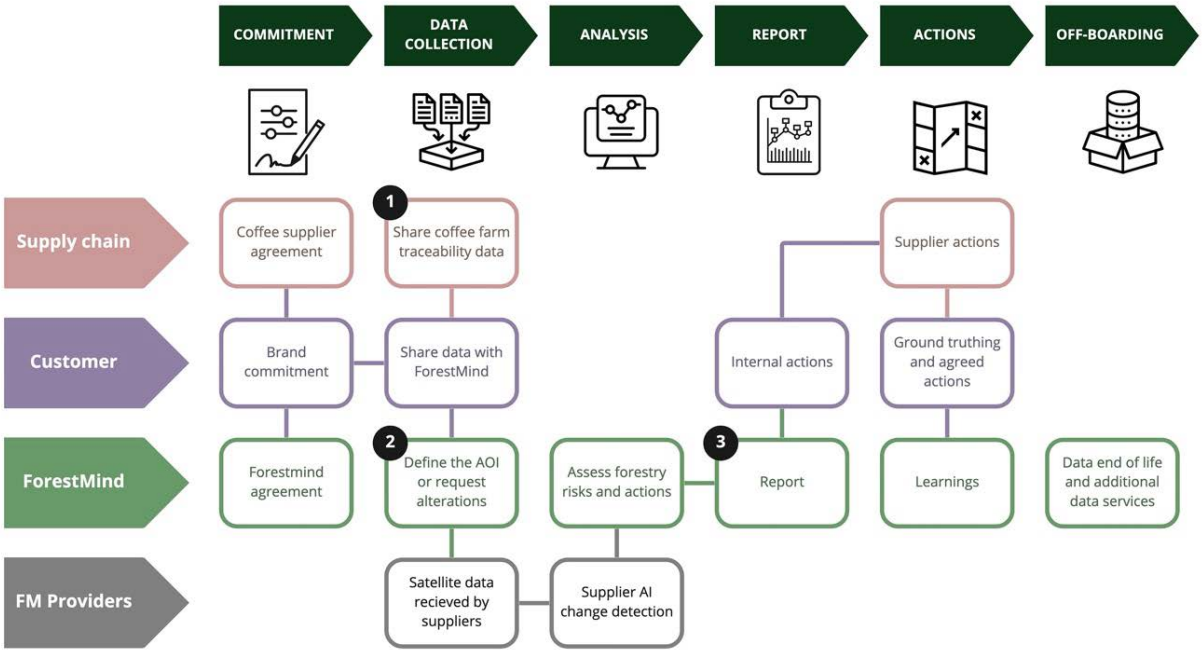



*Figure 2: A service blueprint highlighting the flows of data involving different stakeholders from the start of the service on the left and the end at the right.*

The data flows through interactions in several steps in the provision of ForestMind's services:

## 3.1 Commitment

ForestMind discusses data from the outset of the customer relationship. It is key to understanding what data will be provided by the coffee brand, such as location, points, or field boundaries. The level of traceability data and constraints of imagery available for the geography both have a bearing on the extent and quality of intelligence that can be provided. As part of agreements, the coffee brand includes a code of conduct for suppliers, which includes expectations around environmental impacts such as deforestation.

## 3.2 Data Collection

ForestMind's activities rest on two types of data:

| 1. TRACEABILITY DATA (SUPPLIED BY SUPPLY CHAIN TO DEFINE THE AOI) | | | | | | | |
|---|---|---|---|---|---|---|---|
| Data field | # | No. ASSOCIATE | ADDRESS / TOWN | VARIETIES | SOIL TYPE | x | y |
| Description | Line number in the document | Number given to the farmer by the association | town name | coffee plan variety | SOIL TYPE | Location data | Location data |
| Uses | Key to idenify farm, without using farmers name | Second minimise mistakes attributing deforestation identifyer to | used to verify the approximate farm location of the | Work could be done on climate risk according to crop types | Work could be done on climate risk according to soil types | Locates farm | locates farm |

- **Asset data tracing the supply chain to the production area to defining an Area of Interest (AOI)** – The association's extension workers who support the farmers with agronomical advice will likely survey the farms. The data quality may vary depending on the extension workers' capabilities. Currently, only the latitude and longitude coordinates of the farm's location are useful. There are examples of location data errors, where it is easy to see the locations and address data is inconsistent with the country or region. MRV providers can easily request amendments in these cases, but it is more challenging to identify if there are no obvious inconsistencies, especially where the addressing systems are not mature. ForestMind requested field boundary data, but land rights data in Guatemala are contentious, non-exhaustive and will consequently take time to build trust in accuracy and data sharing. The Sourcing Manager at the coffee brand visits the region to meet suppliers annually. They can verify data or encourage new practices at this point, but it is only practical to visit a small sample of the 263 farms.

| 2. GEOSPATIAL DATA (USED TO MONITOR THE | | | | | |
|---|---|---|---|---|---|
| Data field | Sentinel image | IBat conservation areas | Elevation | Topography | ??? survey data?? |
| Description | 10m multispectral images | Geographically protected areas | Location hightabove sea level | Contours of the land in the geographical area | location of rivers |
| Uses | Used to detect deforestation using machine learning | Understanding the proximity of farms to protected areas | Elevation affects coffee's susceptibility to global heating | Supports intepretation of deforestation data | Location is a risk factor for a range of environmental impacts |

- **Monitoring of the AOI for changes in canopy cover using satellite data** - Satellite images[5] comes from European Space Agency's Sentinel 2 Satellite with a image resolution of 10m. Although it would be hard to identify foliage by eye, computational algorithms developed using machine learning make use of multiple light spectral bands to classify forest and non-forest areas.

## 3.3 Analysis

Once the AOI is defined, ForestMind's analyst requests data from its partner, University of Leicester (UoL), which use a proprietary machine learning algorithm to detect deforestation in satellite imagery. According to ForestMind's accuracy study (see annex 1) UoL's deforestation detection is 80-83% accurate (more recent developments have increased the accuracy in UoL's own studies). While alerts generated from the data are likely to result from actual forest loss, it's still challenging to definitively attribute causes to deforestation events. Deforestation could be caused by farm expansion but also by new farmers or landslides, which are common in coffee regions. At the customer's request, ForestMind can provide farms' geospatial attributes hypothesised to indicate higher-risk suppliers, such as proximity to protected conservation areas and waterways and the farm's elevation (see data ethic concern #3).

[5] *https://www.esa.int/Applications/Observing_the_Earth/Copernicus/Sentinel-2/Introducing_Sentinel-2*

## 3.4 Report

A data report is created from the canopy cover conversion (change detection) data (Annex 2), and the ForestMind Analyst writes a customer-facing document which profiles the farms, highlights important intelligence, and lists recommended actions.

There are usually two types of recommended actions from reports:

1. Steps to improve the supply chain's data maturity, which in this case means verifying or 'ground-truthing' the farm locations and requesting field boundary data. A number of apps are already built for task of creating field boundary data by, for example, asking a farmer to walk the perimeter of their land parcel and then submitting the field boundary data.
2. Corrective actions to ensure the organisation maintains its social and environmental responsibilities.
    a. Ground truth the deforestation event and the cause
    b. Take appropriate actions with suppliers to correct the environmental impact (this will

## 3.5 Actions:

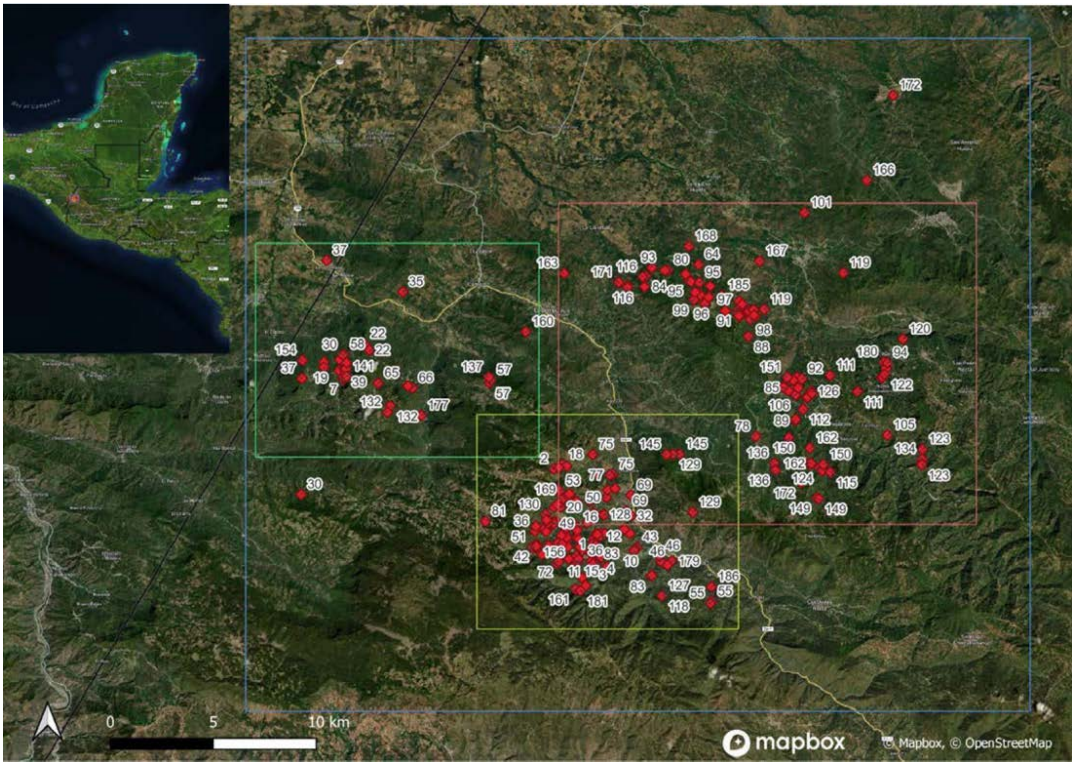
There are currently no industry best practices for responding to deforestation. However, once the events are verified, ForestMind's customers have different options depending on the extent of deforestation. Pending the terms of the regulations, if deforestation breaches the 0.5 ha limit, it may be illegal for coffee to enter UK supply chains.

# 4. Data Ethics Concerns

Data ethics considerations of providing ForestMind's services were explored through three workshops. The table in Annex 2 shows the outcomes from desktop research and analysis using the consequence scanning tool with insights suggesting how the data practices directly in ForestMind's control cascade through the supply chain. The three top concerns identified are described in the sections below:

## 4.1 Reporting only on the coffee farm might indirectly result in deforestation caused by subsistence or other crops

The coffee pilot saw ForestMind monitor 263 given coffee farm locations across a broad geographic area of Guatemala (see the map below). While the clustered farms suggest communities, there are also large amounts of land not monitored by ForestMind between the farm boundaries.




ForestMind aims to report on each farm's area and a buffer zone around the farm area to identify contiguous deforestation. Where farm boundary data is available, it is used; otherwise, a 400m-by-400m box, centred on the farm's reported location, is used to estimate the farm area. The pilot service created a baseline of historic deforestation from the legislation's legal cut-off date (2020) to the current day. Land cleared before December 2020 is considered 'clean' land and can be used to grow commodities for UK export. If farmers want to increase or change their productive area for coffee, they might produce coffee on land used for subsistence or other non-coffee crops before the 2020 cut-off date and move crops not for export onto newly deforested land[6]. Therefore, the impact on net deforestation contiguous to the farm, or even the wider region, could be less than hoped.

[6] *Eliza Zhunusova et al, Forest Policy and Economics: Potential impacts of the proposed EU regulation on deforestation-free supply chains on smallholders, indigenous peoples, and local communities in producer countries outside the EU, Elsevier, (2022), p5*

Coffee brand's codes of conduct and the associations, where applicable, should attempt to mitigate the effect described above with wholistic policies for the farmland. For example, policies require some farmland to be protected and used to compensate for the impact of farming practices. However, the resolve of MRV customers could be tested by deforestation either contiguous to the declared farm area or deforestation caused by subsistence crops further outside the area currently reported by ForestMind.

**Recommendation:** MRV services might ideally detect deforestation at three to four levels of detail:

1. If field boundary data is available to distinguish and monitor changes in the:
    a. declared farming area (current best practice)
    b. a buffer zone around the farm (with research to be carried out on the size of a suitable buffer area) (current best practice)
2. Relate the farms and regions connected to specific farmer associations to indicate the effectiveness of the association's policies for deforestation and agronomical support.
3. Monitoring the extent of deforestation at different geographical scales, such as townships and municipality levels, could give more context to the broader influences of the supply chain and the safest regions to source from. Although it might be difficult to associate deforestation happening in the broader community with the supply chain, if thriving areas with low net deforestation exist, it would be more convincing to demonstrate deforestation free supply chains.

**Limitations:** Unintended consequences of monitoring the coffee farm, and influencing deforestation outside the coffee farm, needs more evidence. However, it is reasonable to be concerned where deforestation rates around the farm are greater than the deforestation in it, ForestMind and customers should be careful about claiming a positive impact on deforestation.

## 4.2 The balance of the intelligence and recommendations encouraging support or penalisation of farmers

Once deforestation is observed and the cause verified within the supply chain, brands must decide how to respond. Following MRV service recommendations, the customer's first action should be to ask the farmers to halt any deforestation activity and negotiate a support package of corrective measures. Given the low education and literacy levels, it may be that the farmer was not aware they were breaking the code of conduct, or social (e.g. crime) or environmental (e.g. climate) pressures may mean the farmer is not receiving sufficient support to meet their subsistence needs.

The alternative corrective measures would be to exclude farmers from the supply chain. This might have serious social outcomes and the opposite unintended effect on deforestation. If the farmer cannot find a buyer at the same premium price, they may be forced to sell to markets with no social or environmental requirements at a lower price. This might encourage the farmer to expand their farm to compensate for the lower income, resulting in deforestation[7].

[7] *Eliza Zhunusova et al, Forest Policy and Economics: Potential impacts of the proposed EU regulation on deforestation-free supply chains on smallholders, indigenous peoples, and local communities in producer countries outside the EU, Elsevier, (2022), p4*

However, if the coffee brand is only made aware of the deforestation after breaching the defined 0.5ha threshold, the coffee could be banned from entering the UK market.

**Recommendation:** ForestMind should seek to offer intelligence on deforestation before it breaches the 0.5 ha threshold. Early warnings from, for example, 0.25 ha deforestation would allow customers to proactively engage associations and farmers to warn them before the farmers are excluded from the supply chain for legal compliance.

ForestMind should also seek to support the development of best practice monitoring response frameworks for MRV intelligence. Industry groups such as Forest Positive Coalition and Speciality Coffee Association would be vital to increasing consistency in the sector.

**Limitations:** Profit margins in speciality coffee are relatively thin across the supply chain, and frequent monitoring might need to be more affordable for smaller brands. However, studies of the patterns, speed, and seasonality of deforestation associated with smallholder farms are still in the early stages. The optimum ratio of report frequency for timely warnings and cost needs to be discovered before the feasible best practices and their consequences can be ascertained.

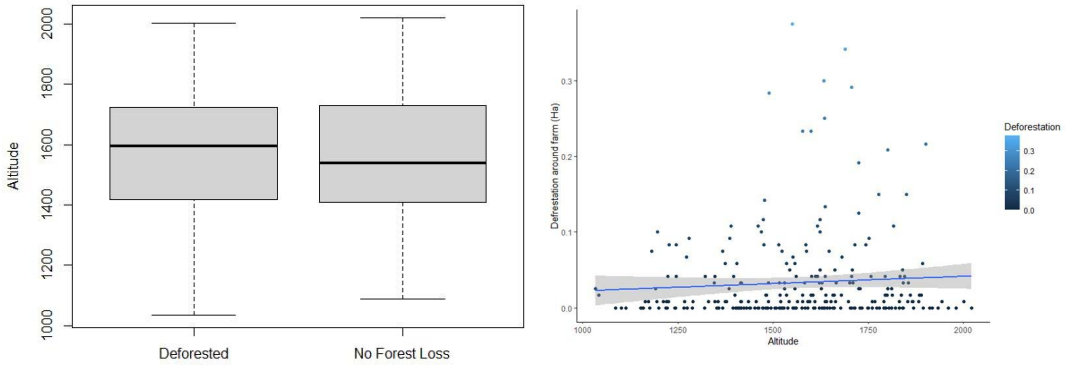
## 4.3 The link between elevation and deforestation risk

Proactively identifying and managing the risks of deforestation would be better than reacting to its consequences and would also enable greater security of supply[8]. The climate crisis means that the optimum growing conditions for coffee are shifting towards higher elevations, with lower elevations starting to see declining yields and lower-quality crops. As yields decline, farmers may need to expand their farms or move them further up the mountainside, causing deforestation in both instances. Coffee brands have shown interest in the elevation of each farm, as well as the distance to water (risk caused by - agri-chemical runoff – not deforestation) and slope (risk of erosion, landslide and soil matrix stability).

| | Farm ID | location | UniqueID | Altitude | Slope- % | Distance To Water (m) |
|---|---|---|---|---|---|---|
| 0 | UC-FID-1 | 1-91.8684708, 15.519464] | UC-FID-1_1461 | 1461 | 18.6465 | 843.261 |
| 1 | UC-FID-1 | 1-91.8755384, 15.5141488] | UC-FID-1_1707 | 1707 | 40.2467 | 1335.31 |
| 2 | UC-FID-2 | 1-91.8858894, 15.5535675] | UC-FID-2 2022 | 2022 | 57.6277 | 1623.63 |
| 3 | UC-FID-2 | 1-91.8835827, 15.5542248] | UC-FID-2.1912 | 1912 | 33.5367 | 1388.2 |

[8] *https://www.climate.gov/news-features/climate-and/climate-coffee*

A rudimentary study was undertaken to determine whether deforestation and elevation are linked. No illegal deforestation was found in the coffee supplychain, but there were a number of instances of canopy changes below illegal levels. Results in the two tables (below) show that canopy cover change are marginally correlated to farms at the higher elevation.




The distribution of canopy chances and no-canopy changes around farms farms is similar (left), but farms with canopy changes in or nearby the given location tend to be at a higher elevation; there is a weak but statistically significant correlation between altitude and canopy changes that might lead to illegal deforestation (right).

The data suggest that higher elevations are at the greatest risk of deforestation, possibly because of increased demand for land in better growing conditions. However, there is little indication that farms at lower elevations are trying to expand their farms to compensate for declining production. Possible explanations of the data patterns could include:

- Locations at the lower elevations, in the most built-up area, could be incorrectly attributed to farm locations.
- If the locations at lower elevations are indeed correct, they might be in built-up areas with limited space for deforestation.
- Deforestation at higher elevations might be new farmers creating farms contiguous to the existing supplier's farm but not in the brand's supply chain.

Field boundary data could support more confident evidence of patterns between deforestation attributable to suppliers and the farm's elevation. However, if MRV services detect deforestation caused by new neighbouring farms while farms at lower elevations decline, there is a risk to supply chain security or coffee grown on deforested land fraudulently entering the supply chain.

**Recommendation:** MRV service providers should be cautious about sending data to clients that associate the farm's location and elevation unless there is a clear understanding of how the data will be used and evidence guiding it. It is recommended that ForestMind continues working with the coffee brand to develop a programme for gathering field boundary data and verifying the locations.

**Limitation:** With the current capabilities, it is unclear how to determine the risk of fraudulently coffee grown on deforested land entering the supply chain.

# 5. Conclusion

The incoming legislation will provide a key driver for ForestMind's business and guiding principles for delivering ForestMind services and the actions for actors in the supply chains. While the legislation has good intentions, research into the effect on smallholders suggests the legislation could have a wide range of unintended negative social and environmental consequences. However, it is in the power of intelligence providers and supply chain actors to ensure that best practices are developed for the legislation's responsible implementation. This non-exhaustive review of data ethics prioritised three key considerations for ForestMind's practices (further considerations in annex 2), with the following actions recommended:

1. In addition to detecting changes in the canopy cover in and around coffee farms, ForestMind should consider detecting deforestation in the broader region that the supply chain might influence. More work is needed on how this can be done responsibly and what practical actions can be taken.
2. ForestMind should attempt to influence customer policies to support farmers rather than exclude them when deforestation is reported. ForestMind can support smallholders by studying the frequency of reporting needed to detect deforestation before it crosses the legal threshold, balanced with cost variables that might affect smallholder inclusion in supply chains.
3. ForestMind should encourage caution if suppliers aim to determine a farm's deforestation risk by associating the available farm locations with elevation and other factors.

# 6. Annex 1 - QA Assessment of Forestmind Forest loss products.

This section provides a quality assessment of the data from Optical Deforestation Detection product provided by the University of Leicester. In the absence of ground data for the region, a visual assessment has been carried out using VHR Planet-NICFI data[9], as well as an existing annual dataset from a third-party deforestation detection open data provider.

## 6.1 Approach to testing accuracy

For each dataset 100 points were randomly selected from two classes: those classified as having forest loss and those not having been classified as experiencing forest loss. The three datasets are:

- Optical Deforestation Detection – University of Leicester (2019-02-07 – 2021-02-22)
- Third-party deforestation detection open data provider (2017-2020)

Each dataset was visually inspected using the Planet-NICFI data, looking at a time series of imagery from before and after the detected forest loss date. Statistics generated are commission (change detected when no change occurred) and omission error (change not detected when change occurred), as well as the percentage accuracy.

## 6.2 Results T2 (at the end of the pilot)

Throughout the project, both the Deforestation Detection datasets have been being worked on to attempt to improve the accuracy metrics. In August 2022, an updated accuracy review has been carried out. The datasets have been tested in two different regions one in Brazil and one in Guatemala, these two sites have been assessed separately for accuracy as they are both geographically different. The region of Brazil is known for growth of Soy and the area of Guatemala for Coffee.

The datasets in this second assessment were:

- Optical Deforestation Detection – University of Leicester (January 2020 – August 2022) (UoL have since made further developments which their own tests could suggest an accuracy of 92.8%[10])
- Third-party deforestation detection open data - (January 2020-August 2022)

| Dataset 2020 | Overall Accuracy | Rate Commission | Rate Omission |
|---|---|---|---|
| Optical Deforestation Detection - Soy Brazil | 83% | 18% | 1% |
| Optical Deforestation Detection – Coffee Guatemala | 80% | 21% | 3% |
| Third-party deforestation detection open data – Soy Brazil | 44% | 57% | 7% |

9 *Planet Team (2017). Planet Application Program Interface: In Space for Life on Earth. San Francisco, CA. https://api.planet.com*

10 *Ivan Reading et al, Due diligence for deforestation-free supply chains with Copernicus Sentinel-2 imagery, University of Leicester, 2023*

# 7. Annex 2 – Sample Deforestation detected

| FL Opt (ha) | Optical Deforestation Map |
|---|---|
| 0.124932 |  |
| 0.0916168 |  |
| 0.083288 |  |

# 8. Annex 3 – Consequence Scanning Work

| # | Within ForestMinds Direct Control | Within ForestMinds influence | Monitoring indirect impacts | Impact (inc. risk and probability level) | Evidence Base |
|---|---|---|---|---|---|
| #1 | Reporting only on the coffee farm, might indirectly result in deforestation caused by subsistence or other crops > | The extent of impacts that customers are willing to take responsibility for > | Legislation guidelines > | Increased deforestation outside the AOI (high concern, high probability) | Evidence comes from peer reviewed research |
| #2 | Based on insights recommend appropriate actions for the level of concern > | Client actions on code of conduct > | Excluded farmers could be forced to sell to alternative sources at lower prices with less ethics > | Increased poverty, and incentives for deforestation (high concern, high probability) | Operations evidenced in conversation with customers, impact evidenced by peer reviewed paper |
| #3 | Associating deforestation risk with attributes that do not have a real correlation > | Bias in sourcing policy > | bias towards larger exploitative farm practices > | Possible worker exploitation and loss indigenous people farming, increases in monocultures (high concern, medium probability) | Data study found no correlation between deforestation and elevation, impact influenced by peer reviewed paper |
| #4 | The lack of industry accuracy standards in deforestation detection makes it hard to compare across providers and regions > | Organisations might choose to purchase MRV services that allow them to make preferential claims > | Lost credibility of remote sensing sector and tolerance of deforestation inconsistent > | Incentivises covering up deforestations and loss of credibility for services (High concern, medium probability) | Performance is observed, and impact is speculated |
| #5 | Confidence in the intelligence, means ground truthing is advised > | Customer has a limited ability to visit site, in the absence of being on the ground customer can call association or pay a contractor > | Relying on 3rd party ground truthing opens the door to corruption, bribery, and blackmail in a country perceived as one of the worlds most corrupt > | Full impacts unknown (Concern low, probability medium) | Performance is observed, impact is speculated based on the perceived corruption index |
| #6 | Risk of service costs exclude small holders from supply chains if brand pass the costs or burden of admin on> | Excluded small holders might be forced to sell to buyers who will only pay a lower price but will accept coffee causing deforestation > | Excluded small holders might be forced to sell to buyers who will only pay a lower price but will accept coffee causing deforestation > | Exposing farmers to possible corruption practices (concern medium, probability low) | Suggested by peer reviewed research on impacts of deforestation regulation, but implications for ForestMind are unknown |
| #7 | Unknown consent levels > | Do the customers really understand data laws and the technology? > | Do the farmers really understand what they are giving consent to? And how does it influence trust? > | Farmers may not like their (concern low, probability low) | Speculative risk based on lack of confirmation |


Satellite Applications Catapult
Electron Building
Fermi Avenue
Harwell Didcot
Oxfordshire
OX11 0QR

Tel: +44 (0) 1235 567999
Email: info@sa.catapult.org.uk