AI-empowered traceability for forest-connected supply chains

> *Effective implementation of regulations such as the EUDR, CITES, and the Lacey Act demands credible, cost-effective methods for verifying harvest locations.*

Deforestation is increasingly recognized as a critical barrier to achieving internationally agreed goals on combating biodiversity loss, mitigating anthropogenic climate change, and achieving sustainable and equitable development. In tropical forests, deforestation is driven largely by the expansion of industrial-scale commodity agriculture, as well as logging for timber. Many countries have responded with 'demand-side' laws requiring due diligence which necessitates knowledge of the harvest origin[1] of such commodities in international trade, most notably the recent European Union Deforestation Regulation (EUDR). Effective implementation of the EUDR, and agreements like the Convention on International Trade in Endangered Species (CITES), or enforcement of the US Lacey Act and relevant sanctions, requires credible and cost-effective methods for scrutinizing declared location of harvest and determining the origin of particular commodities and products in trade.

In a recent peer-reviewed paper published in *Nature Plants*, World Forest ID presents an innovative Artificial Intelligence (AI)-enabled 'origin model', which allows for efficient determination of the origin of agricultural and timber specimens in international trade, based on scaled, ground-truthed data. While the paper demonstrates the effectiveness of the model using temperate tree data, it was developed to facilitate objective verification of geolocation claims and therefore the traceability systems required to comply with policies aiming to tackle the trade in 'conflict', or illegally harvested timber, as well as agricultural commodities originating from illegally converted forests.

This briefing summarizes the implications of the research for a policy-focused audience.

THE INSIGHT SERIES

World Forest ID's Insight series is designed to communicate the outputs of our long form research in a timely manner, by summarizing data snapshots and interim learning. All research is ultimately published in appropriate peer reviewed journals and citations should reflect full articles wherever possible.

Access full article: A framework for tracing timber following the Ukraine invasion

1. The World Trade Organization (WTO) defines 'country of origin' as "*the country where goods were wholly produced or were subjected to sufficient processing*". However, in the context of our work, and throughout this report, we refer to 'origin' as the location of harvest of tree and plant materials used in the manufacture of traded products.

1. Scaled data is actionable data

The World Forest ID origin model published in *Nature Plants* was based on an ongoing initiative which has, to date, collected 8000 ground-truthed tree samples covering seven highly traded timber genera across 11 countries in Central and Eastern Europe. As a result, users can query a reference set that represents a significant proportion of the commercial extraction range of the species in question, in order to determine the most likely point of harvest of an unknown traded sample.

Scale of this sort is only economically viable if multiple laboratories and stakeholders can use the model, which depends on replicability between laboratories. In order to facilitate this, the reference data was developed using non-proprietary protocols and was subject to a 5% replicability test.

> ***Scale of this sort is only economically viable if multiple laboratories and stakeholders can use the model.***

The scale of sampling feeding the model means that results have been actionable in a number of critical real-world scenarios; thereby demonstrating the importance of balancing method innovation with scale to achieve impact. For example, since making our reference model available to real-world users, it has supported numerous investigations into laundering of 'conflict timber' most likely harvested in Russia and Belarus that has entered the EU in the form of plywood and sawn wood products processed in Asia.

In line with our institutional mission to create scaled and actionable data sets, World Forest ID continues to expand the range and granularity of our reference sample data, increasing scope in partnership with national forest agencies, responsible companies, and civil society, to empower decision makers and regulators to protect forests.

2. Combining scientific techniques in one model increases confidence

The peer-reviewed model takes 22 ground-truthed chemical values from 929 geolocated reference samples, extracted from standing trees across the commercial range of the species in question. Each chemical value is the result of analysis using two different scientific techniques; Stable Isotope Ratio Analysis (SIRA)[2] and Trace Element Analysis (TEA)[3]. The full set of measurements is then combined in a state-of-the-art AI process to infer chemical signals for unsampled locations, for example those in conflict zones.

SIRA and TEA are often represented as alternative or competing techniques for establishing product origin, however, by comparing and combining reference data from both, the study demonstrates that optimal results are achieved by fusing multiple data types into a single reference model.

2. Isotopes are forms of the same chemical element with the same number of protons but different number of neutrons and therefore different relative atomic mass.

3. Trace elements are chemical elements that are present in the environment in very small (trace) amounts.

3. Machine learning brings down cost and increases range

The EUDR, which entered into force in 2023, raised the bar for traceability, requiring companies to declare sub-national geolocation of harvest 'plot' for all regulated commodities prior to import, including timber. While the requirement is ambitious, ensuring that its impact is meaningful depends on regulators having objective ways to scrutinize those declarations and, where false, challenge them. There is an extensive body of evidence showing that the effectiveness of global timber trade controls has been undermined by 'laundering' of illegal products, using documents purporting to show harvest in legal and/or certified sustainable areas. Mainstreaming accurate scientific testing for geolocation verification will be critical to ensure that the effectiveness of the EUDR, and similar potential regulations in other markets, is not compromised in the same way.

> *Using AI to infer chemical signals in unsampled locations from available physical reference data.*

While geolocated physical samples are the foundation of our approach, sampling the number of locations necessary to achieve comprehensive spatial results is resource intensive. In addition, ideal collection coordinates can correspond with areas that are in political conflict or otherwise unsafe for sample collectors. Using AI to infer chemical signals from physical reference data, enriched with a wealth of freely available environmental and climatic datasets, can significantly increase the scope and range of reference data, at reduced marginal cost.

4. Spatial modeling facilitates data fusion and increases granularity

World Forest ID's AI-supported geospatial approach enhances the power of chemical analysis to verify harvest locations, integrating environmental variability that is unaccounted for by traditional statistical methods. By interpolating chemical signals across geographies, the accuracy of the model can be improved significantly by holistic consideration of available geospatial forest datasets, for example species ranges, topography, or active logging roads, all of which can be derived from earth observation.

> *The accuracy of the model can be improved significantly by holistic consideration of available geospatial forest datasets.*

The approach captures spatial variability by employing precise GPS data for reference locations, creating a landscape-reflective 'origin model', allowing it to effectively infer chemical profiles in unsampled areas. In this way, it transcends traditional statistical models and arbitrary political boundaries. The sub-national granularity achieved from this approach supports both broad and specific regulatory requirements, including compliance with the US Lacey Act as well as more granular supply chain integrity initiatives. Critically, the approach increases the ability of users to scrutinize harvest claims that are close to critical spatial boundaries, such as protected areas and national borders.

Ultimately, conceptualizing chemical signals as spatial data allows our model to pinpoint harvest geolocation and assess the risk of supply chain laundering more precisely, utilizing both chemical data and, for example, freely available information on species ranges, indigenous territories, deforestation frontiers, and independent monitoring of illegal activity in forestland.

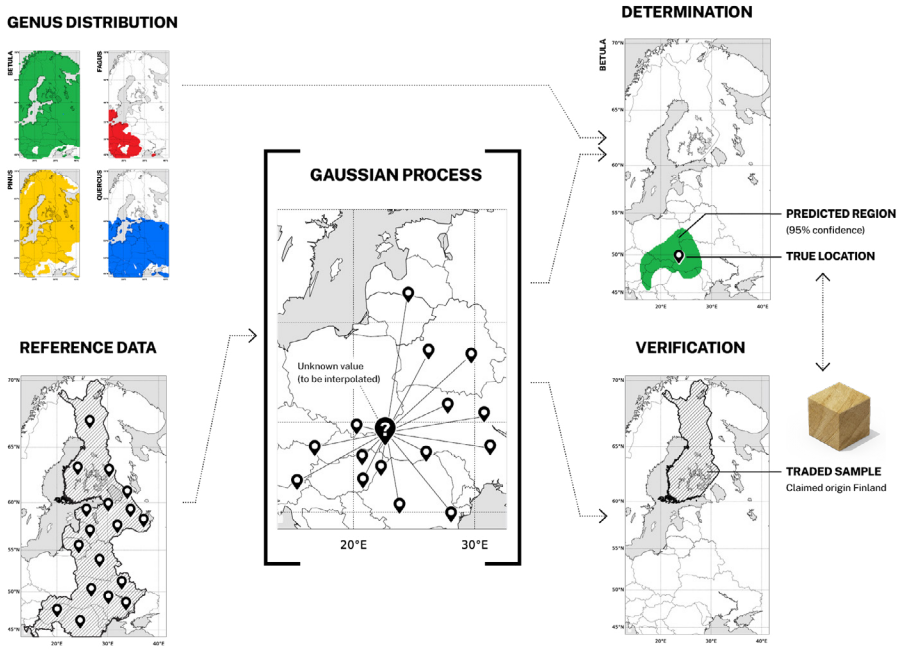
5. Real world application (birch/*Betula* spp.)

The functionality and flexible calibration of the model application for *Betula* spp. (birch) profiled in the paper has been developed in cooperation with users from multiple agencies responsible for enforcing existing timber import regulations, as well as private sector compliance teams seeking innovative ways to verify the claimed origin of material in their supply chains.

A key variable in the model, requested by users, allows for the calibration of results to a preferred level of confidence. In practice, this means that a preference for a 95% accurate spatial representation of possible origin identifies a larger potential area of harvest than an equivalent query seeking a 75% confidence range. In addition to the confidence range, the model can identify a single pixel representing the 'most likely' location of harvest, based on all available data. This unique flexibility brings utility to a wide range of users, from business regulators working in a 'due diligence' context to US Lacey Act prosecutors seeking the most robust results possible. It also allows companies and certification schemes to review the claims of their suppliers in multiple ways, depending on their regulatory and reputational drivers.

The adaptability of our model allows it to serve various stakeholders from regulatory enforcers to private compliance teams, making scaled impact a real-world possibility. By developing a tailored approach for scrutinizing the origin of timber and other forest risk commodities, World Forest ID is setting a new benchmark in ensuring transparent and accountable supply chains.

FIGURE 1: *World Forest ID AI 'origin' modeling*



ACKNOWLEDGEMENTS:

Expert review

Charles (Chip) Barber – Director, Natural Resources Governance & Policy, World Resources Institute

John Simeone – Principal, Simeone Consulting