Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Global patterns, supply-chain flows, and driving factors of forest loss footprints from 2001 to 2022

Mimi Gong

gongmimi@bnu.edu.cn

Faculty of Geographical Science Beijing Normal University

Zhouyi Liu

Ye Li

Xinkun Wang

Shen Qu

Yinglan A

Beijing Normal University

Jin Wu

Advanced Interdisciplinary Institute of Satellite Applications, Faculty of Geographical Science, Beijing Normal University

Guoqiang Wang

Beijing Normal University https://orcid.org/0000-0002-4479-7889

Qiao Wang

Faculty of Geographical Science, Beijing Normal University

Article

Keywords: forest loss footprint, Input-output analysis, interpretable machine learning, deforestation

Posted Date: March 4th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-8944281/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: There is **NO** Competing Interest.

Global patterns, supply-chain flows, and driving factors of forest loss footprints from 2001 to 2022

Zhouyi Liu[1,2#], Ye Li[1,2#], Xinkun Wang[1,2], Mimi Gong[3,4*], Shen Qu[1,2*], Yinglan A[3,4], Jing Wu[3,4], Guoqiang Wang[3,4], Qiao Wang[3,4]

1. School of Management, Beijing Institute of Technology, Beijing, China

2. Center for Energy & Environmental Policy Research, Beijing Institute of Technology, Beijing, China

3. Advanced Interdisciplinary Institute of Satellite Applications, Beijing Normal University, China

4. Faculty of Geographical Science, Beijing Normal University, China

Email: gongmimi@bnu.edu.cn

* These authors are the corresponding authors.

# These authors are co-first authors.

Abstract

Monitoring forest loss is critical to keeping global warming below 1.5 °C, yet achieving zero deforestation remains elusive. Despite extensive national monitoring and high-resolution earth observation data, deforestation persists due to a limited understanding of the human and economic mechanisms driving forest conversion. Here, we develop a state-of-the-art analytical framework to disentangle these dynamics by tracing forest loss—driven by logging, permanent agriculture, commodity production, and infrastructure expansion—through global supply chains to the consumer side, termed the forest loss footprint. This approach bridges the gap between fine-scale remote sensing data and aggregated economic accounts. We further apply machine learning techniques to identify key socioeconomic drivers behind consumption-induced forest loss. Our results show that 38.15% (13.12 Mha in 2022) of global forest loss is embodied in international trade, with five countries (Brazil, the United States, China, Indonesia, and Russia) accounting for 48.16% of global deforestation. China and Russia exhibit polarizing trends that amplify global forest pressures. Among drivers, the shares of food imports and the urban population emerge as dominant factors explaining cross-national patterns of forest loss. This integrated framework offers a robust, data-driven foundation for advancing global forest conservation and promoting equitable responsibility across international supply chains.

Keywords: forest loss footprint, Input-output analysis, interpretable machine learning, deforestation

1. Introduction

Forests, as critical ecosystems on Earth[1], cover 30% of the Earth's land surface[2]. The total value of the world's forests is as much as $150 trillion[3], not only accounting for 1% of global jobs, employing more than 33 million people[4], but also maintaining 80% of terrestrial biodiversity, providing irreplaceable habitats for countless species[5]. More critically, forests serve as one of Earth's largest carbon sinks[6], absorbing about -3.5 ± 0.4 (where − denotes a net carbon gain) petagrams of carbon annually (PgC per year) in the 2010s[7], playing a vital role in climate change mitigation. Despite numerous protection commitments by the international community, global forest cover is estimated at –10.2 million ha per year in 2015–2020[8], particularly in ecologically valuable mangrove forests and tropical rainforests[9,10]. These losses are alarming in the face of worldwide climate targets in controlling global temperature within 1.5°[11].

Global deforestation, a persistent environmental challenge, has undergone significant changes in patterns and intensity over recent decades. Latest data reveals that global forest loss reached 8.1 Mha in 2024, equivalent to losing 15 football fields of forest every minute. The geographic distribution of deforestation is highly uneven, showing distinct regional concentration: Brazil and Indonesia alone account for nearly half (47%) of tropical forest deforestation. Brazil alone accounts for one-third of global tropical forest deforestation, losing approximately 1.7 Mha annually, while Indonesia accounts for 14% [5,12]. Regarding driving factors, between 2001 and 2024, approximately 34% of global tree cover loss resulted from permanent land use change, while in tropical primary rainforest regions, this proportion reached 61%[5]. From an economic perspective, deforestation is broadly categorized into four main types: commodity-driven permanent agriculture, forestry activities (logging), and the expansion of settlements and infrastructure. These data highlight deforestation as

58 an urgent global challenge that requires immediate attention [9,13-15].

59 To address the deforestation crisis, a multi-level global governance system has emerged.
60 Numerous forest protection policies and initiatives have been implemented worldwide, demonstrating
61 remarkable diversity, including regional zero-deforestation commitments (such as the Amazon Soy
62 Moratorium)[16], consumer country-led regulations (the EU Deforestation Regulation, EUDR)[17],
63 industry sustainability certifications (RSPO palm oil certification)[18], and corporate zero-deforestation
64 pledges (the New York Declaration)[19]. These policy instruments can be categorized into four main
65 types: market incentive mechanisms, regulatory constraints, voluntary standards, and multilateral
66 agreements [20-23], reflecting the global society's multifaceted efforts to address deforestation. However,
67 previous research indicates that geographic spatial constraints and complex market dynamics
68 significantly limit the effectiveness of these policies. For instance, while the Amazon Soy Moratorium
69 reduced local deforestation rates by 80%, it led to a 25% increase in deforestation activities in the
70 Cerrado savanna region[16]; similarly, palm oil certification programs in Kalimantan have caused
71 deforestation activities to shift to other parts of Borneo, with an average displacement rate of 40%[17].
72 This "leakage effect" severely undermines the overall effectiveness of these policies, highlighting the
73 importance of cross-regional coordination[16,17,19].

74 The inefficiency of policy implementation arises from a combination of market, governance, and
75 institutional constraints. First, due to the lack of coordinated planning and cross-border responsibility-
76 sharing mechanisms, forest protection policies in one region often lead to deforestation activities
77 being displaced to neighboring areas with weaker regulations[24,25,16]; Second, insufficient transparency
78 within indirect supply chains exacerbates the problem—Zu et al. found that more than 60% of
79 agricultural products are traded through indirect channels, making sustainability verification and

accountability difficult[26]; Third, regulatory enforcement remains weak in many tropical forest countries, where limited governance capacity, high enforcement costs, and the vast spatial extent of forested areas constrain effective monitoring and control[25,27]. Collectively, these intertwined challenges highlight the critical need for systematic tracing of forest loss chains from consumers to producers to ensure accountability and strengthen global deforestation governance.

In recent years, academia has developed various methods to track responsibility chains for global forest loss. Hoang et al. pioneered the analysis of deforestation footprints, combining land-use change with international trade data to quantify consumer countries' "embedded responsibility" in tropical deforestation for the first time, finding that 15-18% of tropical deforestation is directly related to international trade [28]; Zu et al. through detailed supply chain mapping, discovered that indirect procurement accounts for 12-100% of agricultural commodity chains with deforestation risk averaging 40% higher[26]; Hoang et al. extended this analytical framework to conservation policy domains, using spatial modeling to analyze overlaps and conflicts between agricultural production and high conservation value areas, identifying that 22% of global biodiversity hotspots face direct threats from agricultural expansion[29]. Although these studies provide important insights into understanding global mechanisms driving forest loss, they are primarily limited to single-dimension analyses—either focusing on spatial patterns, specific agricultural products, or supply chain structures—lacking an integrated comprehensive framework that organically combines economic activity types, spatial flows, and driving mechanisms, as well as in-depth analysis from phenomenological description to mechanistic interpretation.

Moreover, these efforts to link global economic activity to forest dynamics often rely on environmentally-extended multi-regional input-output (EE-MRIO) models[30,31], which often source

forest-related data from statistical agencies such as the FAO. However, these top-down allocations of national statistics face several limitations. Firstly, they typically capture only aggregate flows (e.g., total timber harvest) and net stocks (e.g., forest area), without distinguishing among the diverse drivers of forest loss [24,32-35]. Second, cross-country comparisons are often inconsistent because national datasets differ in definitions, reporting procedures, and temporal intervals[36-40]. Third, gaps in temporal coverage are frequently filled through interpolation, introducing modeling uncertainties[41-44]. By contrast, remote sensing datasets provide consistent, high-resolution observations of forest dynamics with global coverage over time, but their spatially explicit structure is not readily compatible with aggregated economic frameworks such as MRIO models [45-47]. Supplementary Excel 1 summarizes the major global MRIO databases and their corresponding forest data sources. This methodological divide between detailed physical observations and aggregated economic systems underscores the need for a bottom-up, empirically grounded approach to attribute forest loss events to their proximate drivers with high spatiotemporal accuracy—an essential step toward evidence-based policymaking and shared accountability in forest conservation.

This study traces global forest loss from 2000 to 2022, linking high-resolution remote sensing data with international trade to quantify and interpret the footprints of forest loss driven by four major land conversion types: permanent agriculture, logging, hard commodities, and infrastructure expansion. By connecting production and consumption flows, we map the spatial–temporal dynamics of forest loss embodied in global supply chains and classify countries according to polarizing and equalizing trends in forest resource use to inform targeted policy interventions. Furthermore, we employ state-of-the-art machine learning techniques to identify the socioeconomic driving factors underlying consumer-induced forest loss across different land conversion types. Specifically, this

study addresses three key questions: (1) What are the spatial–temporal patterns and flows of forest loss footprints associated with global supply chains from 2000 to 2022? (2) How can countries be typified based on their temporal trajectories of forest resource use? and (3) Which socioeconomic factors explain cross-national variations in forest loss footprints?

2. Results

Our analysis first links the newest remote sensing forest data with activities that drive the conversion of forest ecosystems to the global trade of products. Forest loss footprint represents the extent of forest loss driven by four activities to meet a given country's final consumption: hard commodities (HC), logging (LOG), permanent agriculture (PA), and settlements and infrastructure (SI). This footprint can be partitioned into domestic and foreign forest-loss footprints. For instance, a forest-rich country like Brazil, which hosts vast forest resources, has a forest loss footprint that reflects its consumption-driven capacity to drive domestic and foreign forest loss, especially from hard commodities and logging activities on the production side. In contrast, a country like Iraq, with negligible forest coverage, has a forest loss footprint that drives only foreign forest loss, as its consumption leads to the conversion of foreign forests for logging or commodities embodied in the supply chain, reflecting a consumption perspective. Our results primarily focus on the forest loss footprints of 189 consumer countries, sourced from 201 countries where forest loss occurs, and examine how these four land-conversion activities contribute to these footprints from 2001 to 2022. We then categorize countries based on their spatio-temporal flow dynamic and apply machine learning (ML) techniques to understand the driving factors.

Figure 1 illustrates the structural composition of forest loss footprints across four major land conversion drivers from 2001 to 2022. International trade accounts for approximately 38.15% of the total global forest loss footprint, while domestic consumption represents 61.85%. Within international flows, trade between similar income-level countries constitutes the most significant component (16.46%), followed by low-to-high income flows (14.96%) and high-to-low-income flows (6.73%). This aggregate pattern masks substantial variations across activities. PA, HC, and LOG demonstrate high levels of globalization, with significant spatial decoupling between consumption and environmental impact. In contrast, settlement and infrastructure (SI) remains predominantly localized, with domestic activities driving 86-90% of its footprint. This divergence—globalized trade-driven deforestation versus localized domestic development—highlights the need for differentiated policy approaches: supply chain governance for internationally traded products and domestic land-use planning for infrastructure development.

While international flows for LOG are concentrated among countries of the same income level (20-27%), it also exhibits the highest proportion of flows driven by lower-income countries among the four activities. Lower-income countries import raw logs to meet their domestic processing demand, while regional processing networks operate among countries of similar income levels. The predominant international flows of PA and HC to high-income nations (18% and 28%, respectively) clearly delineate the ecosystem impacts of global consumption inequality, in which high-income consumer demand is displaced through global supply chains, ultimately manifesting as forest loss in low- and middle-income countries.

In terms of absolute scale, demand from consumer countries drives the most forest loss from PA at 5.2–9.23 Mha, followed by Logging (3.45–6.95 Mha). HC and SI constitute a secondary

category that operates on a smaller scale. Over 2000-2022, the total footprint from HC shows a clear upward trend, indicating its growing importance. In contrast, LOG and PA exhibit relative stability in total volume, while SI shows a more cyclical pattern. A notable temporal pattern emerges: the highest values for all drivers occurred in 2017, while the lowest, except for PA (2001), were recorded in 2003.

While the long-term trend of forest loss appears relatively stable, the total value exhibits significant inter-annual volatility. Notably, between 2012 and 2017, land-conversion activities exhibited sharp, synchronized fluctuations. For instance, the PA activity peaks around 2012 and 2016–2017 were closely mirrored by those of HC. This lockstep pattern suggests they are responding in unison, not to diffuse consumer demands, but to unified price signals from integrated global commodity markets. Despite the domestic driver, SI, exhibited a slower, multi-year cyclical pattern, with distinct five-year declines from 2007–2011 (from 0.24 Mha to 0.17 Mha) and again from 2017–2021(from 0.25 Mha to 0.17 Mha), distinguishing the high-frequency volatility of globally-exposed sectors from the slower cycles of domestic policy and investment.

Although all four drivers contribute to forest loss, their scale and structural dynamics differ markedly. LOG provides a stable baseline of impact, while its sensitivity to global commodity trade heightens the volatility of PA and HC. The pronounced cyclical peaks observed in PA and the cyclical declines in SI may reflect the influence of global economic cycles or policy shifts on forest loss.

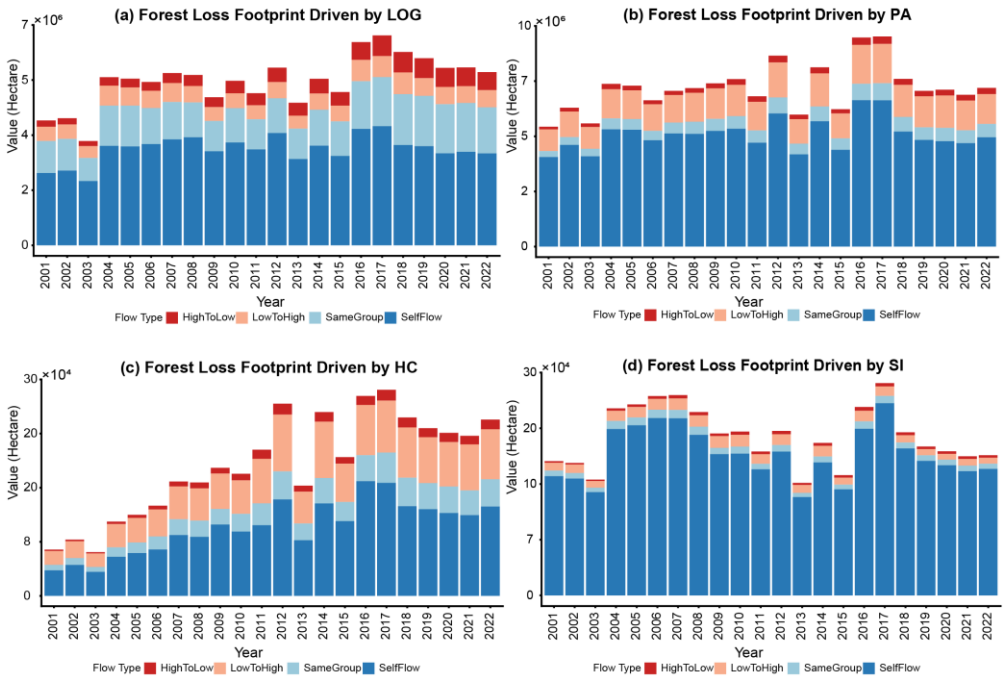


Figure 1: Temporal distribution of forest loss footprint by land-conversion activities and income flow categories (2001-2022). The figure illustrates annual forest loss footprint (hectares) driven by four major land-conversion drivers: logging (LOG)， permanent agriculture (PA), hard commodities (HC), and Settlement& infrastructure (SI). Each stacked bar represents the total annual forest loss footprint driven by the specific economic activity, with color-coded components indicating different cross-border consumption flows: red segments show economic activity in higher-income countries driven by lower-income countries' consumption (HighToLow); orange segments represent economic activity in lower-income countries driven by higher-income countries' consumption (LowToHigh); light blue segments indicate forest loss driven by countries within the same income group (SameGroup); and dark blue segments represent domestically driven forest loss (SelfFlow).

Figure 2 further reveals the structural evolution of forest loss footprints from 2001 to 2022, illustrating the flow patterns for the total footprint (Figure 2a) and the footprint driven by PA (Figure 2b).

While domestic consumption remains a significant component of the global forest loss footprint, international trade has profoundly reshaped its structure between 2001 and 2022, driving a transformation from a unidirectional "South-North" model to a complex multi-directional network.

The shift in global trade flows was characterized mainly by domestic consumption and intra-group exchanges among peer-income economies (Fig. 2a). The most prominent feature across all periods is the prevalence of thick, grey-colored flows, which represent the vast majority—consistently over 60%—of the footprint being consumed within the producer country or traded between economies of the same income level. Key consumers such as the USA, Brazil, Russia, and China collectively accounted for a substantial share of the total global footprint (51.33% in 2001 and 48.16% in 2022), yet satisfied most of their demand for embodied forest loss internally (**SI-Table 1**). Global forest loss footprint in 2001-2022 is further displayed in **SI-Excel 2**.

Against this backdrop of high internal consumption, the transformation at the international level becomes even more significant. In 2001, global flows were primarily defined by flows from low- to higher-income (South-North) countries, which accounted for 46.71% of all international flows. By 2022, while this "South-North" flow remained significant at 39.22%, the network had evolved into a more intricate landscape, with "North-South" flows (from high-to-low-income countries) nearly doubling their share from 10.60% (2001) to 17.65% (2022). The most representative example of this shift is the markedly increased outflow of footprint from the USA and Canada to China, which grew from 1.46% (2001) to 11.40% (2022) of the two countries' outflows. Deforestation accountability becomes increasingly complex as 'Southern' consumption now drives 'Northern' land-use change, effectively displacing environmental burdens.

Major countries exhibit distinct role differentiation in the global forest loss footprint network. Brazil and Russia consolidated their positions as primary international suppliers, with footprint outflows constituting 15.26% and 12.07% of all international flows, respectively, in 2022. China evolved from a major footprint producer to a core hub for global footprint inflows, with its share surging from 3.1% in 2001 to 9.7% in 2022. The USA presents a contrasting pattern. While possessing one of the world's largest total footprints (1.21 Mha), its role is defined by massive domestic consumption rather than international trade. Its contribution to global trade was modest, with outflows accounting for just 4.40% of total international flows, indicating America primarily satisfies demand through internal resources.

Distinct dynamics within key land-conversion activities drive this overall transformation of the global footprint. Focusing first on PA (Figure 2b), we observe that **PA's international footprint has undergone a profound structural transformation, evolving from a highly concentrated "South-North" model into a complex, dual-core system.** This evolution was driven by concurrent diversification among producers and a fundamental reorientation among consumers. Initially, production outflows were overwhelmingly dominated by Brazil, which accounted for 30.83% of total international outflows in 2001, with Southeast Asia as the other significant producer region (29.35%). However, the producer countries progressively diversified, with more emerging in different parts of Latin America and, most notably, across Africa's multi-centric and decentralized network (e.g., Mozambique and Angola), which was already visible by 2011 and became a defining feature by 2022, with their combined share of outflows growing from 0.20 Mha in 2001 to 0.49 Mha in 2022. This expansion transformed the producer network, even as Brazil maintained its role as the single largest producer, still responsible for 31.02% of global international outflows in 2022.

The convergence of producer-side trends and parallel transformation on the consumer side established the network's dual-core structure. In the early 2000s, global consumption footprints were dominated by developed countries (the Global North high-income economies), with the top 10 OECD countries particularly prominent, collectively accounting for 30.36% of the total global footprint. However, the most significant change over this period was the rise of emerging economies, epitomized by China, as major consumption centers. China's share of global consumption footprint surged from 1.38% in 2001 to 4.82% in 2022, transforming it into a "many-to-one" convergent receiving hub. This development, driven by substantial growth in flows from Brazil, Africa, and Southeast Asia to China, did not replace traditional "South-North" flows—flows from developing to developed countries still accounted for a significant share of total inflows in 2022. Instead, it established a powerful parallel "South-South" axis—flows from developing countries to emerging economies—creating a dual-core global system that has reshaped the dynamics of agriculture-driven land conversion.

The footprint flows of HC and LOG also demonstrate a profound structural evolution (**SI-Figure 1-2**). **The consumer structure for HC has clearly transitioned from a US-centric, single-pole model in 2001 (USA: 26.81%, CHN: 5.81% of total footprint) to a stable "dual-core" structure dominated by China in 2022 (USA: 10.00%, CHN: 17.87%).** Accompanying this process, international flows have shifted from being unilaterally dominated by low-to-high income transfers to a multi-directional, balanced network in which both high-to-low and low-to-high flows coexist. On the supply side, producers of HC are more diverse, with regions such as Latin America, Russia, Southeast Asia, and Africa collectively supporting the global supply network. In contrast, the producers for LOG are more concentrated in traditional resource-rich nations such as Russia and

Brazil, featuring a significant growth in Russia's footprint driven by the consumption of China (from 0.01% to 17.18% of Russia's footprint) and the dual role of Canada as a high-income country that is also a major producer. This structure mirrors the prevailing global log trade dynamics: forest-abundant developed nations such as Canada supply raw timber to Asia's manufacturing hubs for value-added processing. Such patterns exemplify the resource-to-processing supply chains that connect high-income exporters with rapidly industrializing economies, driving the pronounced high-to-low income flows evident in **Figure 1**, which are predominantly fueled by surging demand from China and Southeast Asia.

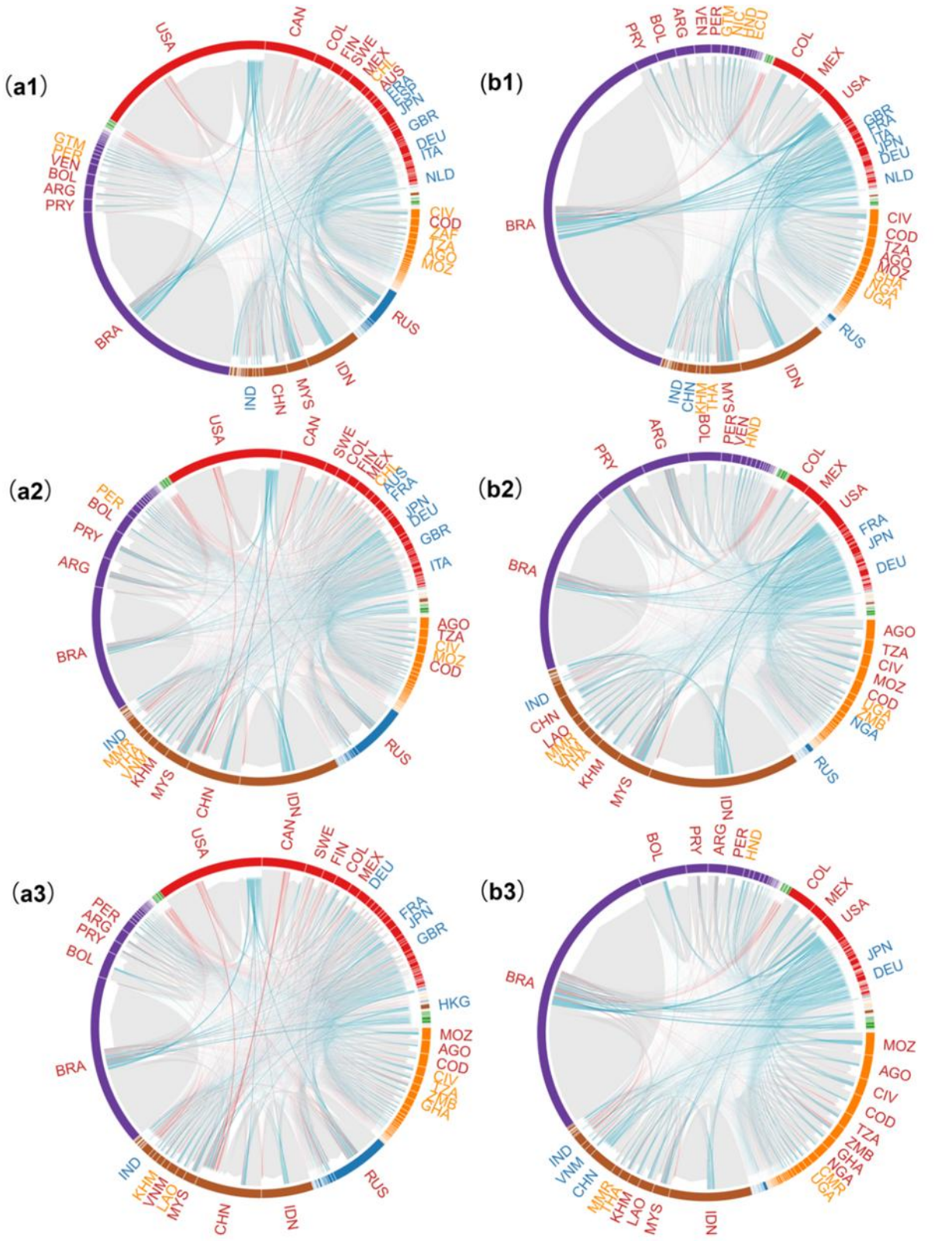

(a1)
(b1)
(a2)
(b2)
(a3)
(b3)
Flow color
Same Income
High to Lower Income
Low to Higher Income
Ring color
E_Euro_Asia
MENA
OECD
Africa
LAC
E_S_Asia
Oceania
Label color
Producers Top 25 only
Consumers Top 25 only
Both Top 25

Figure 2: Country-to-country forest loss footprint flows from 2001 to 2022 across income groups and regions. (a) Total forest loss footprint. (b) Forest loss footprint driven by Permanent Agriculture (PA) at the country level for 2001, 2011, and 2022. The ring colors around the circle represent countries' geographic regions, with countries arranged according to their regional position. Flow colors indicate income-related patterns: gray flows represent forest loss driven by countries of similar income levels; red flows show forest loss footprints in higher-income countries driven by lower-income economies; and blue flows indicate footprints in lower-income countries driven by higher-income economies' consumption. The visualization maintains the same regional color coding while highlighting how PA contributes to cross-border impacts on forest loss. Countries are labeled according to their significance in forest loss networks: red indicates countries ranking in both the top 25 consumers and producers of forest loss footprint, orange highlights countries among the top 25 producers only, and blue designates countries among the top 25 consumers only. Regional classifications include Eastern Europe and Asia (E_Euro_Asia), the Middle East and North Africa (MENA), OECD countries, Africa, Latin America and the Caribbean (LAC), East and South Asia (E_S_Asia), and Oceania, as shown in the legend.

Figure 3 reveals a dual pattern of polarization and equilibration in global forest loss outsourcing over the 2001–2022 period.

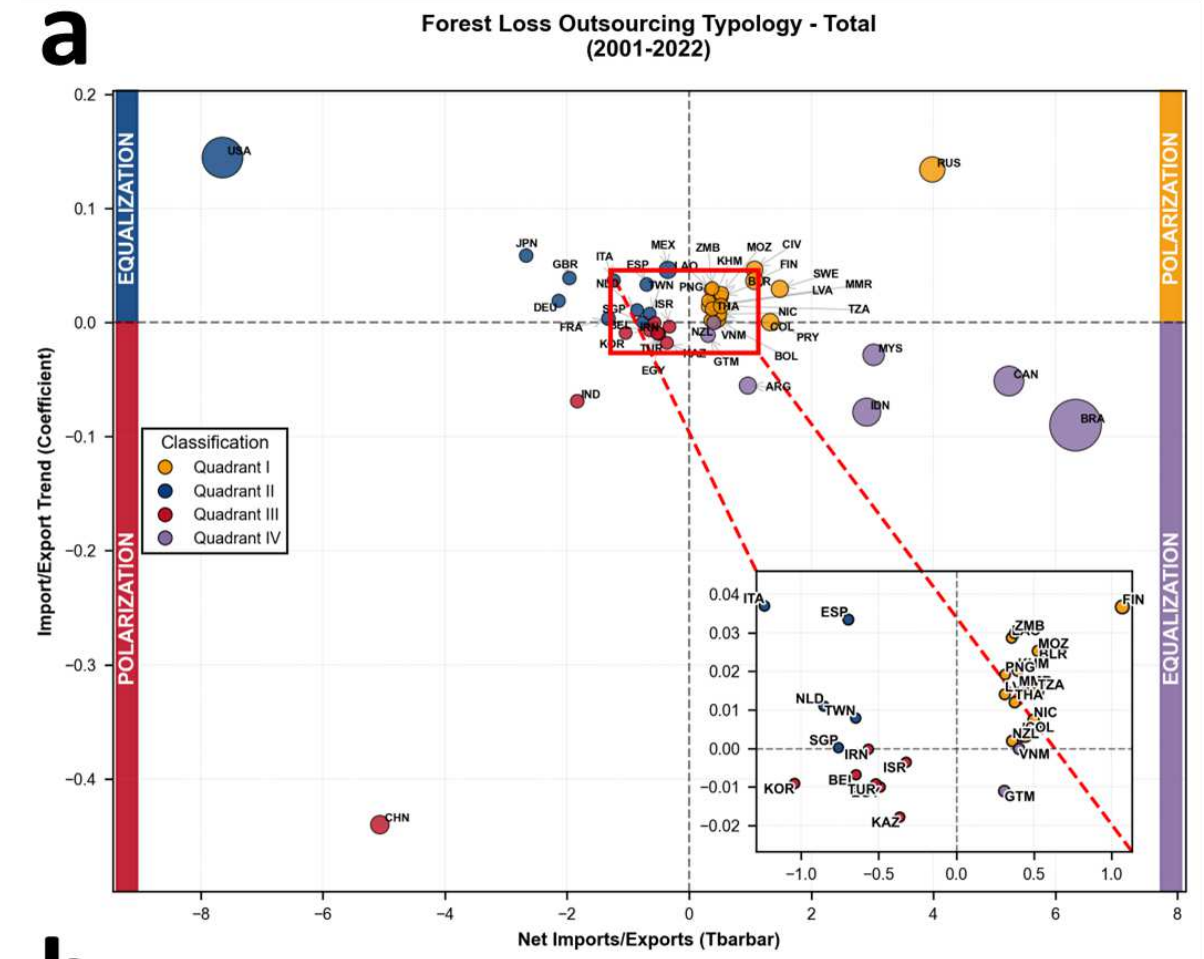
Quadrants I (orange) and III (red) (Figure 3a) reflect polarization trend**.** Countries in Quadrant I, such as Russia (RUS), are net producers with intensifying export trends, reflecting their deepening role as resource suppliers. Conversely, Quadrant III nations, including China and India, exhibit increasing consumer dependence, underscoring their growing reliance on external forest resources (Figure 3b). This bipolar configuration indicates an escalating imbalance in forest resource flows between specific regions.

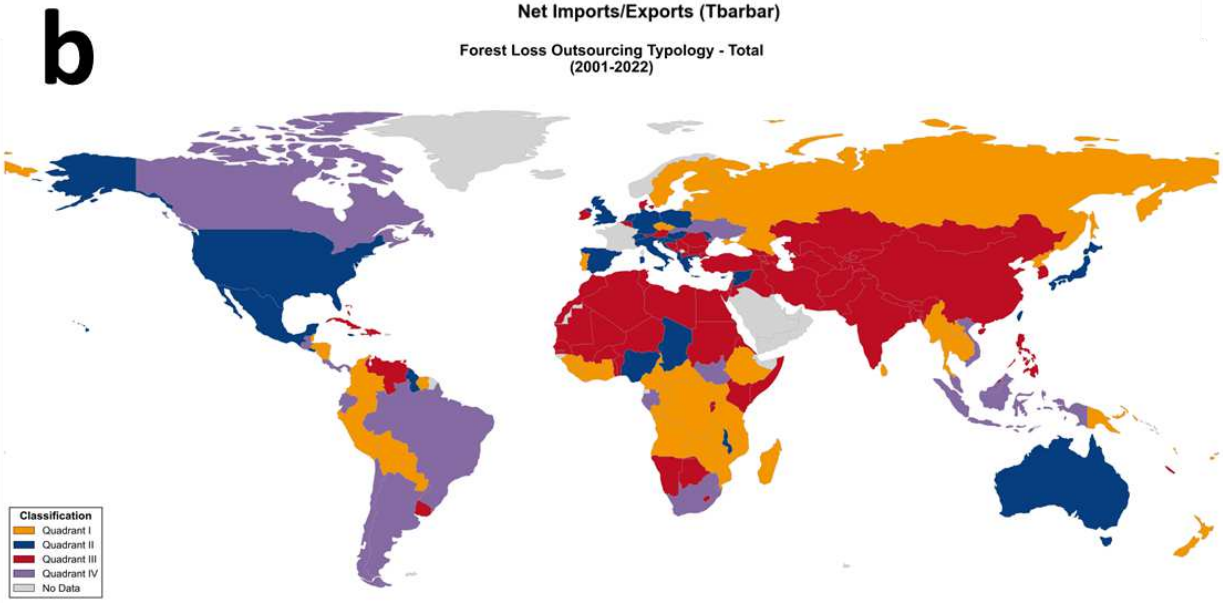
Equilibration tendencies emerge in Quadrants II (blue) and IV (purple). Quadrant II comprises developed economies such as the United States, Japan (JPN), and the United Kingdom (GBR), which, despite being net consumers, show declining import trends—suggesting reduced external forest dependence. Quadrant IV countries, including Brazil, Canada, and Indonesia, are net producers with weakening export trends, potentially reflecting a shift toward more sustainable resource management practices.

Bubble sizes reveal the disproportionate influence of key players. Brazil stands out as the largest net exporter, with outsourcing volumes far exceeding those of other nations; however, its pronounced downward trajectory (Quadrant IV) may signal the effectiveness of Amazon conservation policies. The United States, as the second-largest bubble in Quadrant II, indicates a transition from a major consumer to a more balanced trader. Although China occupies a modest position in the lower-left region, its upward trend (deep red) foreshadows a potential emergence as a dominant net producer in the coming decades.

Our driver-disaggregated analysis reveals that **polarization is predominantly driven by HC and LOG** (**SI-Table 4**). Emerging economies such as China and India consistently occupy Quadrant III (red) across both drivers, indicating intensifying import dependence for timber and mineral-related forest products. Conversely, Russia maintains a Quadrant I (orange) position in hard commodities and logging, reinforcing its role as a primary resource producer. This divergence underscores that **extractive industries are the principal contributors to the bipolarity of resource flows**.

In contrast, **equilibration tendencies are more pronounced in PA and SI drivers**. Brazil, despite being a significant net producer in the total footprint (Quadrant IV, purple), exhibits a shift toward Quadrant I (orange) in HC, suggesting sectoral heterogeneity in its resource management strategies. The United States transitions from Quadrant II (blue) in most drivers to Quadrant III (red) in PA, reflecting its reliance on agricultural imports. Canada demonstrates a notable shift from Quadrant IV (purple) in logging to Quadrant II (blue) in SI, potentially signaling **policy-driven reductions in outsourcing urbanization-related forest loss.** The typology of global forest loss outsourcing and maps for each land conversion activity are displayed in **SI-Figure 3-10**.


a
Forest Loss Outsourcing Typology - Total
(2001-2022)
EQUALIZATION
POLARIZATION
POLARIZATION
EQUALIZATION
Classification
Quadrant I
Quadrant II
Quadrant III
Quadrant IV
Import/Export Trend (Coefficient)
Net Imports/Exports (Tbarbar)
USA
RUS
CHN
IND
BRA
CAN
IDN
MYS
ARG
EGY
GTM
BOL
PRY
NIC
COL
TZA
MMR
LVA
SWE
FIN
CIV
MOZ
ZMB
KHM
MEX
ITA
ESP
GBR
JPN
DEU
FRA
KOR
TUR
ISR
SGP
NLD
TWN
PNG
THA
VNM
NZL
KAZ


b
Forest Loss Outsourcing Typology - Total
(2001-2022)
Classification
Quadrant I
Quadrant II
Quadrant III
Quadrant IV
No Data

326

327

Figure 3: Global forest loss outsourcing typology (2001–2022). a, Four quadrants reflecting net forest loss importers/exporters from 2001-2022 and temporal importing or exporting trends. **b**, Global map displaying the spatial distribution of forest loss outsourcing typology, with colors corresponding to the four quadrants in panel **a**.

The x-axis represents the time-averaged net imports/exports, with positive values indicating net exporters and negative values indicating net importers. The y-axis shows import/export trend coefficients derived from ordinary least squares regression of annual footprint volumes, with positive values indicating intensifying exports or declining imports, and negative values indicating declining exports or intensifying imports. Countries are classified into four quadrants based on their positions: Quadrant I (orange) represents net exporters with increasing trends, Quadrant II (blue) net importers with decreasing trends, Quadrant III (red) net importers with rising trends, and Quadrant IV (purple) net exporters with decreasing trends. Bubble size indicates the absolute magnitude of forest loss outsourcing in 2001. The red-dashed line represents the regression between the net trade position and the temporal trend. The inset provides a zoomed view of countries clustered near the origin.

Figure 4 identifies five features accounting for nearly 60% of model prediction variance to explain forest loss footprint: Food imports (% of merchandise imports) (food_imports_pct) (15.77%), Urban population (% of total population)(urban_pop_pct) (13.70%), Forest area (% of land area)(forest_area_pct) (11.50%), Terrestrial protected areas (% of total land area)(protected_areas_pct) (10.19%), and Manufacturing, value added (% of GDP)(manuf_value_added_pct) (8.67%). The food import percentage emerges as a composite indicator integrating socioeconomic development, trade structure, and environmental policy

effectiveness, with powerful associations with permanent agriculture-driven deforestation. **Its influence intensified progressively from 2008 onward** (**Figure 4b**, panel 1), coinciding with global commodity price volatility and agricultural trade liberalization1—a pattern reflecting how international food security concerns amplify demand for agricultural expansion in exporting nations. Other feature contributions to the forest loss footprint and land conversion activities are described in **SI-Figure 11-20.**

Urbanization explains logging-driven forest loss in consumer countries. In logging-specific scenarios, urban_pop_pct (16.57%) surpasses food_imports_pct (15.91%) as the dominant predictor (Figure 4a, blue curve). High-urbanization countries (red points) concentrate in positive SHAP regions (0 to +0.6), demonstrating that urban expansion is the principal driver to explain logging-driven forest loss in producer countries. **Temporal analysis shows that urban population growth consistently exerts a stable influence on logging-driven forest loss (Figure 4b, panel 2, blue line), in contrast to the fluctuating impacts observed for other drivers.** This stable relationship indicates that urban population expansion has consistently contributed to logging-related forest loss since 2001. Notably, gdp_per_capita and pop_density exhibit markedly elevated importance in logging contexts (~7.7% each), with high-GDP regions (red points) corresponding to positive SHAP values, indicating that economically advanced nations can explain intensified logging pressures and drive forest loss in producer countries.

Protected area coverage in consumer countries shows differential patterns in explaining deforestation. Protected_areas_pct maintains relatively stable explanatory importance for logging-driven forest loss (~7% throughout 2001–2022, Figure 4b, panel 3, blue line), while showing a declining trend for permanent agriculture-driven forest loss after 2010 (green line decreases from

~12.5% to ~11%). The SHAP dependence plot reveals a **non-linear, inverted-U-shaped relationship between protected_areas_pct and the forest loss footprint (SI-Figure 32**). Consumer countries with moderate protected area coverage (approximately 4.5–20%) show the strongest positive associations with forest loss footprint (SHAP values peak around 0.2-0.4). In comparison, both low-coverage (<4.5%) and high-coverage (>20%) countries exhibit weaker or negative associations. This pattern suggests a threshold effect, in which the relationship between consumer country conservation levels and producer country forest loss varies non-monotonically across different ranges of protected area coverage.

Manufacturing value added in consumer countries shows similar importance patterns in explaining the different types of deforestation. In the circular diagram (Figure 4a), manuf_value_added_pct (F4) exhibits moderate importance for logging scenarios. The temporal analysis reveals similar trends across deforestation types (Figure 4b, panel 4). For permanent agriculture-driven forest loss, manuf_value_added_pct shows increasing explanatory power (the green line rises from ~8.5% in 2001 to ~9.5% in 2022). In contrast, for logging-driven forest loss, the importance increases more gradually (the blue line rises from ~5.5% to ~7.5% over the same period). Both trends exhibit upward trajectories, but the rate of increase differs between the two deforestation drivers, suggesting that consumer countries' manufacturing sector characteristics have become increasingly relevant in explaining forest loss patterns, particularly for agricultural expansion. The complete temporal dynamics of feature importance across all forest loss drivers are displayed in **SI-Figure 21**.

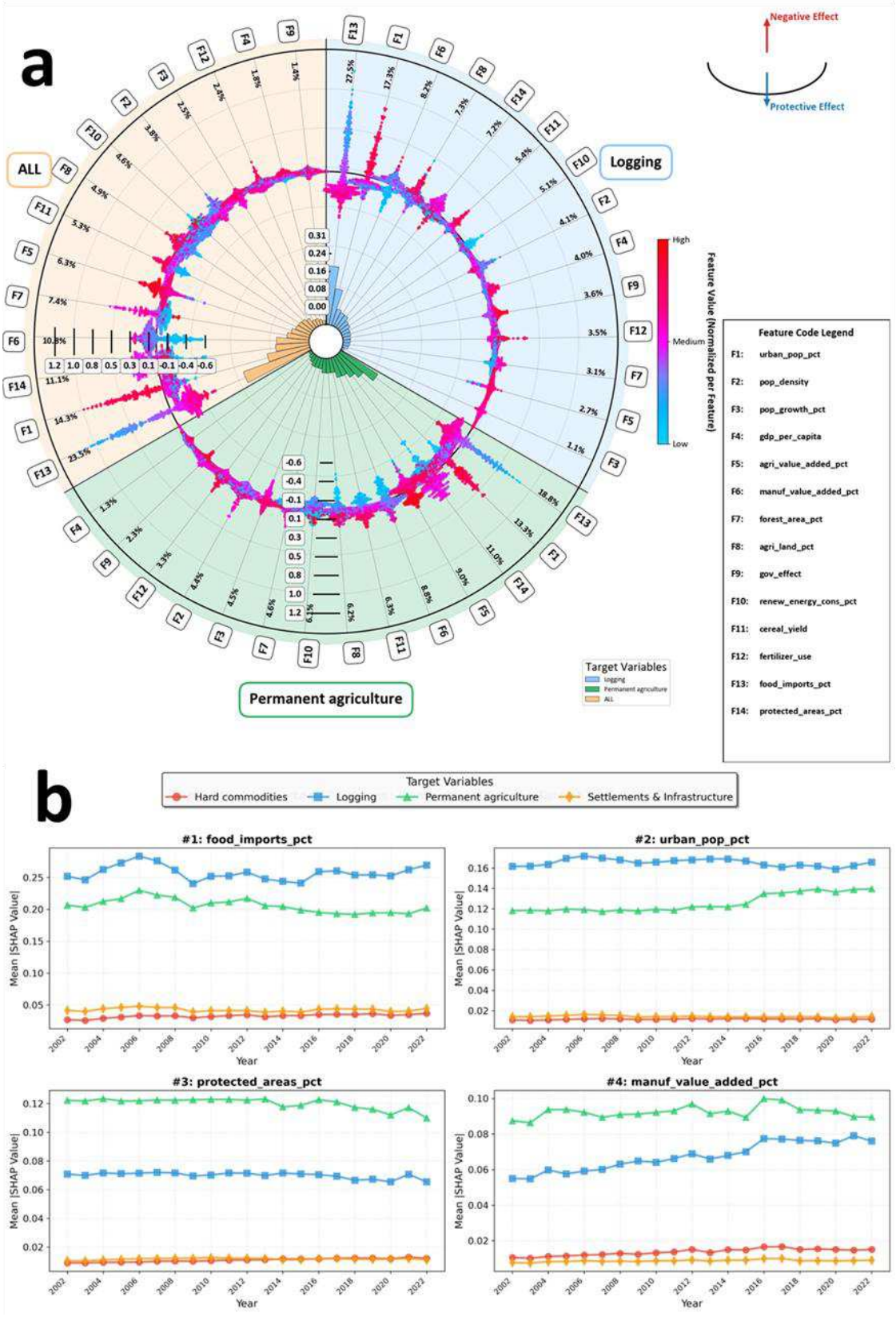


Figure 4: SHAP value visualization showing how consumer country characteristics explain embodied forest loss footprint driven by different land-conversion activities in producer

countries (2001-2022). a, Radial SHAP plot displaying feature contributions (F1-F14) across different forest loss scenarios. The plot shows SHAP values for all forest loss types combined (orange-shaded region), logging-driven forest loss (blue points), and permanent agriculture-driven forest loss (green points and green-shaded region). Each point represents a country-year observation. The radial axis indicates SHAP value magnitude, where deviations toward the periphery (positive SHAP values, 0 to +1.2) indicate features associated with increased forest loss footprint, while deviations toward the center (negative SHAP values, -0.6 to 0) indicate features associated with reduced forest loss footprint. The color gradient from blue (low feature values) to red (high feature values) shows the actual magnitude of each feature for individual observations. Features are arranged circularly (F1-F14), with each feature's contribution displayed as a swarm of points extending radially from the center. **b,** Temporal evolution of mean absolute SHAP values for the top four features (2001–2022) across four land-conversion drivers: hard commodities (red), logging (blue), permanent agriculture (green), and settlements & infrastructure (orange). Each panel shows how feature importance changes over time for explaining different types of forest loss footprint.

3. Discussion

To unravel the coupling between global consumption and deforestation, we develop an integrated framework to quantify forest loss embodied in international trade. We couple high-resolution deforestation data with MRIO modeling to track the footprints of forest loss through global supply chains. Nations are classified by net footprint flow trends, revealing divergent trajectories for tailored governance. We apply XGBoost with SHAP values to assess 14 socioeconomic indicators explaining

consumption-driven forest loss. This approach identifies deforestation patterns from a consumer perspective, providing a scientific foundation for equitable global forest governance.

Our results reveal a bifurcated architecture of global forest loss that fundamentally challenges conventional trade – deforestation narratives. Tradeable commodities — including permanent agriculture, hard commodities, and logging—operate through telecoupled supply chains that spatially decouple production from consumption. In contrast, infrastructure-driven losses follow territorially embedded logics serving domestic needs. This dichotomy underscores the necessity of supranational coordination to govern globalized commodity flows, alongside place-based policies addressing localized land-use pressures.

Crucially, we identify a dual structural narrative: a stable internal loop—domestic consumption accounting for over 60% of total forest loss—forms the systemic baseline, while a dynamic external loop reflects shifting geoeconomic power relations, transitioning from unidirectional South–North exchanges to increasingly complex multipolar networks. The rise of South–South trade axes and intermediary supply chains has diffused environmental responsibility, weakening traditional North–South accountability paradigms and calling for new governance architectures to manage resource appropriation in a multipolar world.

Our typological analysis shows that forest-loss outsourcing follows distinct evolutionary trajectories shaped by both polarization and equilibration dynamics. Polarizing pathways, primarily driven by extractive industries, generate structural asymmetries in the global system. For example, Russia is increasingly expanding its role as a major supplier of forest loss, while China is transitioning to a dominant consumption-driven driver. In contrast, equilibrating pathways observed in many

developed economies often reflect impact displacement rather than genuine ecological decoupling: declining logging footprints are counterbalanced by rising dependencies on imported agricultural commodities. This sectoral heterogeneity—illustrated by Brazil's dual role as an overall mitigating exporter yet an intensifying supplier in hard-commodity (HC) sectors—demonstrates that forest-loss outsourcing is not a single unified process. Instead, it comprises multiple, divergent mechanisms that demand sector-specific rather than one-size-fits-all policy interventions.

Our machine learning analysis uncovers the underlying mechanisms through which socioeconomic drivers shape national forest-loss footprints. Food import dependency exhibits a tripartite differentiation pattern: (i) low import levels indicate responsibility internalization in resource-abundant countries operating under "self-production–self-consumption" regimes; (ii) high import levels in high-income economies reflect strategic outsourcing that exploits global efficiency differentials; and (iii) high import levels in resource-poor nations reflect vulnerability-driven reliance on foreign production. Protected areas show marked sectoral divergence: while effective in curbing logging-related loss, they remain permeable to agricultural expansion, suggesting declining regulatory effectiveness when confronted with profitable land-use competition. Meanwhile, industrialization generates structurally inelastic timber demand, producing a development-pathway lock-in in which manufacturing-based economies systematically perpetuate forest loss through persistent material requirements.

Policy Implications

Our analysis identifies structural patterns in the global footprint of forest loss that necessitate a shift toward differentiated, multi-layered governance architectures. Based on empirical evidence, we propose two complementary policy advisories. First, we urge the development of multi-tier

responsibility-tracking systems, supported by blockchain-verified supply-chain records, to trace forest impacts across complex production networks. Because nearly 40% of forest loss associated with globally traded commodities occurs through telecoupled supply chains, it requires cascading accountability systems that transmit responsibility along the entire production chain. Multinational corporations should be mandated to disclose product-specific forest loss footprints, with contractual mechanisms that assign automatic liability for supplier-level deforestation—for example, when China imports Brazilian soybeans for animal feed and subsequently exports pork to Europe, all participants in the supply chain—producers, processors, and final consumers—would share proportional obligations for forest protection.

Countries should adopt adaptive border adjustment instruments that set tariff rates or sustainability requirements that reflect the exporting country's development status and ecological performance, thereby internalizing cross-border environmental externalities. For example, we recommend establishing South–South cooperation compacts with shared monitoring protocols and sustainability criteria calibrated to national food-security conditions. These partnerships would help correct asymmetries in governance capacity and reduce the displacement of forest pressure across developing economies.

Second, for forest loss driven by domestic infrastructure and locally embedded development, we recommend establishing territorial stewardship bonds that require subnational governments to provide restoration-cost guarantees, thereby incentivizing long-term land stewardship and reducing the risk of irreversible ecological degradation. Moreover, our results reveal that while protected areas effectively deter logging-related forest loss, they remain vulnerable to agricultural encroachment, which leads to forest loss, necessitating dynamic buffer-zone governance. We propose deploying real-

time agricultural-pressure detection systems, including satellite-based monitoring that identifies crop expansion near forest frontiers and triggers immediate regulatory responses. For instance, the appearance of new farmland within protected areas along Amazonian boundaries would automatically trigger enforcement alerts.

Limitations and Future Research

Our study has several limitations and can be improved in three key aspects. First, the analysis is constrained to the national scale to maintain consistency with country-level economic accounts across years. A critical future direction is the development of a fully spatially explicit framework for attributing on-the-ground forest loss to consumers. Such an approach would enable each discrete forest loss event, driven by specific land-conversion activities, to be traced back to its ultimate consumption driver, marking a paradigm shift from macro-level attribution to micro-scale event tracking.

Second, we predicted the four land conversion activities using 14 driving factors with separate XGBoost models. While these models effectively capture temporal dynamics, they only consider a limited set of variables and omit other potential factors, which may introduce bias that could be reduced by incorporating additional variables in future studies.

Finally, our analysis lacks a comparison between our remote sensing-derived loss footprints and forest-related footprints derived from national statistical accounts (e.g., the FAO Global Forest Resources Assessment (FRA) and National Forest Inventory (NFI)). The absence is due to discrepancies between these two types of datasets arising from differing definitions, methodologies, and data collection procedures, making direct comparison impractical [49-53]. Both of them provide critical dimensions of forest resources worldwide. A key research frontier lies in

developing harmonization and calibration frameworks that bridge this gap—for example, through transfer functions or unified accounting standards—to strengthen the compatibility and cross-validation between remote sensing observations and national statistical systems, thereby providing a more coherent baseline for global environmental footprint assessments.

Despite these limitations, our study makes two key contributions: (1) Compared to mainstream forest footprint analysis in existing MRIOs using national statistics in measuring forest stock and flows, we systematically resolve the incompatibility between high-resolution remote sensing data and macroeconomic models. We aggregate pixel-level forest loss events by different land conversion activities into economically traceable flows and link them through the global supply chain. The efforts leverage the most recent global satellite observations of forest loss across time and create an alternative, consistent, and transparent forest loss footprint to complement existing forest-related footprints for nationwide policies across years. (2) We apply state-of-the-art techniques from phenomenon interpretation, archetype identification, and driver explanations to systematically identify the forest loss footprint flows and rank the deep-seated socio-economic driving factors of forest loss in the supply chain. This provides a scientific data-driven basis for designing more targeted policy interventions.

4. Methods

4.1 The calculation of the forest loss footprint

The calculation of the forest loss footprint involves three main steps: 1) summarizing forest loss per country per driver per year, 2) constructing bilateral trade flows per driver per year, and 3) calculating deforestation embodied in trade.

Step 1: Calculate forest loss per driver per country per year.

We use the country statistics directly from Sims et al.'s (2025) data as the baseline because it is one of the most recent forest loss datasets powered by AI models and provides information on drivers at a global scale [48]. Sims's final products include spatially explicit raster data with a 1 km × 1 km resolution per forest driver, covering 191 forest-holding countries from 2001 to 2022. It groups forest loss drivers into seven categories: permanent agriculture, hard commodities, shifting cultivation, logging, settlements and infrastructure, wildfire, and other natural disturbances. We use this product to create spatially explicit maps of forest loss footprints and leverage its country-level statistics to understand the spatiotemporal dynamics of forest loss, thereby linking these findings to steps 2 and 3.

Step 2: Identify forest loss drivers and their linkages with input-output tables.

We then follow the method of Hoang et al. (2021) to link four anthropogenic drivers of forest loss (permanent agriculture, settlements and infrastructure, hard commodities, and logging) to the Multi-Regional Input-Output (MRIO) database. We link the losses from Sims's drivers to the Resolved Exiobase version 3 (REX3), a global MRIO database that divides the global economy into 163 representative sectors, including fisheries, logging, mining, and others, which are directly linked to permanent agriculture, settlements and infrastructure, and hard commodities. The REX3 database

encompasses 189 countries and 163 industry sectors, with time-series data spanning 1995 to 2022. This MRIO database provides comprehensive coverage of agricultural products and has already applied country-specific deflators[49,50], without requiring further deflator adjustments.

Building on this mapping, we consider the definitions of drivers in Sims's calculation, the definitions of REX3's sectors, and the linkages that Hoang et al. (2021) used to build the drivers of deforestation, which they then used to construct linkages with REX3 in the study. For example, the settlements and infrastructure associated with forest loss include the expansion and intensification of roads, settlements, urban areas, and built infrastructure. The hard commodities driver mainly covers mining and energy infrastructure, including mining at all scales, power lines and plants, oil extraction and refineries, wind and solar farms, and flooding from hydroelectric dams.

Step 3: Mapping forest loss footprint

The renowned laureate Wassily Leontief conceived input-output (IO) analysis; today, in accordance with United Nations guidelines, over 100 national statistical agencies routinely publish standardized IO tables that depict flows of goods and services among sectors and final consumers. For example, the transactions matrix T measures the flows between industries, where each element T_{jk} represents the monetary flow of goods or services from sector j for use in sector k. The coverage extends from primary sectors (agriculture, quarrying, mining) to secondary sectors (manufacturing, construction, utilities) and tertiary sectors (trade, hospitality, transport, communications, storage, and services). The set of sectoral forest loss footprint forms the so-called satellite account within the MRIO model. The satellite account contains non-monetary inventories of each sector (e.g., related to forest loss footprint), which can be coupled with the economic MRIO system to quantify how global consumption and international trade drive forest loss.

Gross output for each sector equals intermediate use plus final demand. Leontief's basic accounting identity can be expressed as: $\mathbf{1}^V V' + \mathbf{1}^T T' = T\mathbf{1}^T + Y\mathbf{1}^Y = x$, where $\mathbf{1}$ denotes a summation vector of {1, ...,1}, the prime ($'$) indicates transposition, and x is total output. By defining the technical coefficients matrix as $A = T\hat{x}^{-1}$,where $\hat{x}$ is the diagonalized vector of sectoral outputs, the identity transforms into the well-known Leontief system: $Ax + Y\mathbf{1}^Y = x, \quad (I - A)x = Y\mathbf{1}^Y, \quad x = (I - A)^{-1}Y\mathbf{1}^Y$. Here, I is the identity matrix and $L=(I-A)^{-1}$ is the Leontief inverse. Matrix L quantifies the sectoral output induced by one additional unit of final demand, tracing supply chains from final use back through all intermediate stages and is used in Equation 1.2.

To map the forest loss footprint, we attribute, for every forest loss pixel on the ground, responsibility to each consumer country through the global supply chain. We create spatially explicit pixel-level maps and aggregate them to the country level. The traditional MRIO model is insufficient to depict the complex spatiotemporal patterns of embodied forest loss across regions and drivers. To better identify the primary targets and drivers of forest loss, we link the map of the spatial distribution of forest loss drivers to a MRIO model to explore further the high-resolution forest loss footprints driven by international trade in primary commodities.

Therefore, the pixel-level footprint analysis method calculates forest loss footprint F in country r driven by country s in equation 1.1:

$$F^S = \sum_{hr} D_h^r \frac{\sum_i f_{hi}^r \sum_{tj} L_{ij}^{rt} y_j^{ts}}{\sum_i m_{hi}^r} \tag{1.1}$$

F^s represents a spatially explicit forest loss footprint map driven by consumer countries (1km×1km). The forest loss maps (D) are categorized by forest-holding countries (r) and drivers (h). The embodied forest loss (numerator of the fraction, expressed as fLy, as detailed in equations 1.2-1.4) is normalized by the total forest loss (m) for each driver (h) in forest-holding countries (r) and

each industrial sector (*i*), to downscale values from nations to pixels (the denominator). Both the forest loss maps and the embodied forest loss are expressed in absolute value. The resulting fraction yields the pixel-level attribution of forest loss to consumer countries, which is then aggregated for every production country *r* by driver or product *h*. The forest loss maps (*D*) are prepared in step 1 with the shapefile layers for each driver *h* in each forest-holding country *r* and deforestation intensity *f* for each production sector *i*.

The equations below are to understand *fLy* in equation 1.1. In general, we follow the logic of equation1.2.

$$FLF = F + IM - EX \tag{1.2}$$

Forest loss footprint (*FLF*) is equivalent to the forest loss plus the forest loss embodied in imports minus the forest loss embodied in exports. *F* represents the forest cover loss in the production country, measured and calculated using Sims's driver datasets. *IM* is the loss embodied in imports, and *EX* is the forest loss embodied in exports.

Following the logic in equation 1.2, we first compute the adjusted consumption in equation 1.3 and then calculate the forest loss footprint for each country in equation 1.4.

$$AC = \sum_{tj} L_{ij}^{rt} y_j^{ts} + \sum_{t \neq s,j} L_{ij}^{rt} y_j^{ts} - \sum_{t \neq s,j} L_{ij}^{rs} y_j^{st} \tag{1.2}$$

Adjusted consumption (*AC*)= total consumption + imports - exports. *L* ($\in \mathbb{R}^{(r*i)\times(s*j)}$) is the Leontief inverse, measuring the transformation between origin sector *i* of export country *r* and destination sector *j* in import country *s*; *y* ($\in \mathbb{R}^{(r*i)\times s}$) is the final demand of import country *s* for the sector *i* of country *r*. *t* is both the last sale in the consumption and import term and the country of final consumption in the export term.

We then decompose the total loss in the production country into the embodied deforestation

in consumer country *s* using the MRIO model:

$$CF^s(Production) = \sum_{ri} f_i AC \quad (1.3)$$

Where CF^s is the total deforestation embodied in international trade driven by consumer countries, *f* is the deforestation intensity for each sector *i* (forest loss area from Sims's data divided by gross output). *AC* represents adjusted consumption, which is measured in equation 1.3. The CF^S is expressed as *fLy* in 1.1.

4.2 Systematic Framework for Categorizing Countries Based on Embodied Forest Loss Through International Trade over 2001-2022

We developed a producer-consumer dual-perspective framework to quantify forest loss outsourcing via international trade, integrating multi-regional input-output (MRIO) analysis with time-series regression to classify countries into four quadrants based on their net trade positions (Tbarbar) and temporal trends (coefficient) over 2001-2022 (**SI-Table 3**). This approach enables systematic identification of regions exporting environmental burdens (net exporters with intensifying trends, Quadrant I) versus those reducing external dependencies (net importers with weakening trends, Quadrant II), thereby revealing the dynamic reconfiguration of global forest resource flows driven by consumption patterns rather than production activities alone.

Forest Loss Accounting: Producer vs. Consumer Perspectives

Following Malik et al. (2022), we quantified forest loss from two complementary perspectives to disentangle production-based versus consumption-based responsibilities. The **producer perspective (Q)** attributes forest loss to the territorial region where deforestation physically occurs, reflecting direct resource extraction within sovereign boundaries. The **consumer perspective (F)**

reallocates forest loss to regions consuming products whose supply chains embody deforestation, capturing indirect environmental impacts transmitted through international trade networks. Mathematically, for region *p* and forest type *q* in year *t*, the producer-based footprint $Q_t^{q,p}$ represents territorial forest loss, while the consumer-based footprint $F_t^{q,p}$ aggregates forest loss embodied in all products consumed by region *p*, regardless of production location.

Based on MRIO analysis, we define the forest loss outsourcing matrix as $S^{Q\times R\times T}$, where $S_t^{q,p}=\sum_i Q_{i,t}^{q,p}-\sum_{i,j} F_{j,t}^{q,srp}=Q_t^{q,p}-F_t^{q,p}$, where the matrix is $\bar{S}=\bar{Q}-\bar{F}$,where the matrix is expressed with an overbar notation indicating summation over product categories. Note that the regional indices in the footprint tensor are inverted; in the above summation, *p* is treated as both the production region and the consumption region. If $S_t^{q,p}>0$, this indicates that in year *t*, region *p* is a net exporter of primary forest resource *q*; conversely, if $S_t^{q,p}<0$, then *p* is a net importer.

To analyze temporal evolution, we calculate the time-averaged outsourcing value. We use the temporally averaged forest loss footprint (denoted by double overbar notation), defined as $\bar{\bar{S}}=\bar{\bar{Q}}-\bar{\bar{F}}$, where $\bar{\bar{Q}}^{q,p}=\sum_t Q_t^{q,p}=\sum_t\sum_i Q_t^{q,i,p}$, and $\bar{\bar{F}}^{q,p}=\sum_t F_t^{q,p}=\sum_{i,t} F_{i,t}^{q,p}$, with *N* representing the number of years in the study period.

As described above, the temporal dynamics of global forest loss footprints driven by international trade constitute the core of this work, and the quadrant classifications are as follows:

1. Quadrant I: Net Export with Increasing Export Trend ($\bar{\bar{S}}^{-q,p}>0$ and $\bar{S}^{q,p}>0$)

Regions in this quadrant are (on average) net exporters, and over time, their net exporter status has been strengthening.

2. Quadrant II: Net Import with Decreasing Import Trend ($\bar{\bar{S}}^{-q,p}<0$ and $\bar{S}^{q,p}>0$)

Regions in this quadrant are (on average) net importers, and over time their net importer status

has been weakening (and may even have transitioned to net exporters).

3. Quadrant III: Net Import with Increasing Import Trend ($\bar{\bar{S}}^{-q,p} < 0$ and $\bar{S}^{q,p} < 0$)

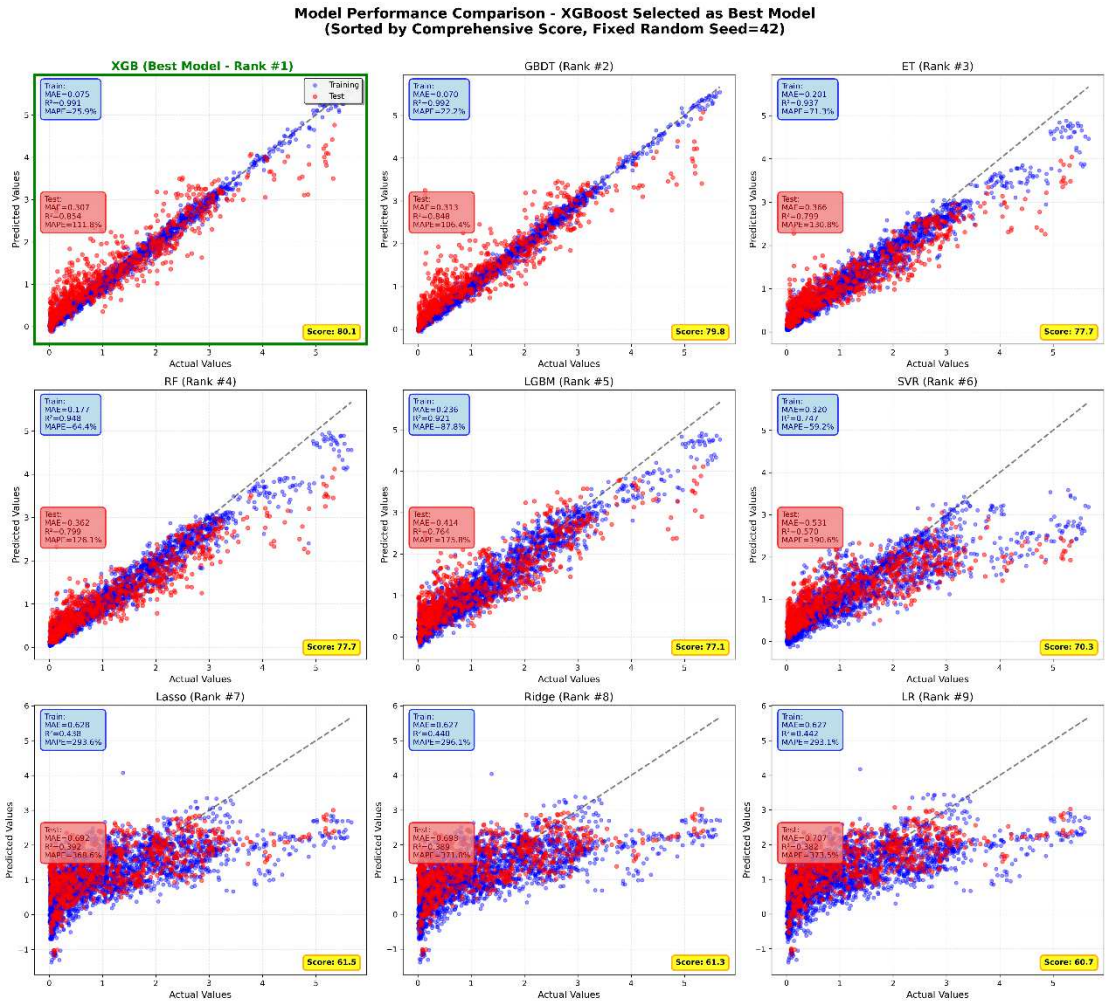
Regions in this quadrant are (on average) net importers, and over time, their net importer status has been intensifying.

4. Quadrant IV: Net Export with Decreasing Export Trend ($\bar{\bar{S}}^{-q,p} > 0$ and $\bar{S}^{q,p} < 0$)

Regions in this quadrant are (on average) net exporters, and over time their net exporter status has been weakening (and may even have transitioned to net importers).

To enhance visual clarity and focus on countries with substantive trade positions in Figure 3, we applied a bilateral filter: |Tbarbar| > 0.3, excluding regions with near-zero time-averaged net trade balances (representing autarkic economies or data quality issues). Bubble sizes in scatter plots were scaled proportionally to export volumes in the first year (2001) to encode baseline trade magnitudes, with size ranging from 30 (base_size) to 500 (max_size) square points.

4.3 Machine learning techniques to explain the forest loss footprint from 2001 to 2022.



Step 1: Data Construction and Feature Engineering

We constructed a comprehensive global dataset spanning 169 countries from 2001 to 2022 to quantify the socioeconomic drivers of forest loss footprint. The dataset integrated 14 core features across four dimensions(**SI-Table 2**): (1) Population and Urbanization (urban population percentage, population density, population growth rate); (2) Macroeconomic Indicators (GDP per capita, agricultural value-added percentage, manufacturing value-added percentage); (3) Land Use Patterns (forest area percentage, agricultural land percentage); (4) Governance and Environment (government effectiveness, protected area coverage, renewable energy consumption, cereal yield, fertilizer use, food imports percentage). Missing values were imputed using k-Nearest Neighbors (k=5) algorithm, which estimates missing entries based on the five most similar observations in the feature space.

We firstly did correlation analysis (**SI-Figure 22**) and numerical features were log-transformed to reduce skewness and stabilize variance across countries with heterogeneous economic scales. The cleaned and transformed dataset contained 3,549 samples, which were used for all subsequent analyses.

Temporal Feature Engineering. To capture nonlinear temporal dynamics and lag effects, we implemented systematic feature engineering: (1) Nonlinear time trends: year_normalized = (year - 2001) / (2022 - 2001) and year_squared = year_normalized squared; (2) Lag features and change rates for critical variables (forest_area_pct, gdp_per_capita, pop_density): feature_lag1 = feature at time t-1, and feature_change = feature at time t minus feature at time t-1; (3) Interaction terms between key drivers and temporal trends: forest_area_pct_x_year, gdp_per_capita_x_year, and gov_effect_x_year.

The final feature set comprised 25 variables (14 base plus 11 engineered), designed to capture cumulative effects, short-term shocks, and dynamic policy responses.

Step 2: Time-Series Data Splitting and Model Selection

Data Splitting. To ensure temporal consistency and prevent data leakage, we employed an 80/20 time-based split: Training Set (2001-2017, n=2,704 samples) and Test Set (2018-2022, n=845 samples). All 169 countries appeared in both sets. For Support Vector Regression (SVR), features were standardized using StandardScaler fitted on the training set.

Model Comparison. We systematically compared nine regression algorithms: Linear Models (Linear Regression, Ridge Regression, Lasso Regression); Support Vector Machine (SVR with RBF, linear, and polynomial kernels); Ensemble Methods (Random Forest, Extra Trees, Gradient Boosting, LightGBM, XGBoost)(**SI-Figure 23-27**). All experiments used fixed random seed (seed=42) and single-threaded execution (n_jobs=1) for reproducibility.

Bayesian Optimization. We employed Bayesian optimization via Optuna framework with Tree-structured Parzen Estimator (TPE) sampling. Each model underwent 50 intelligent trials, minimizing 5-fold cross-validation Mean Absolute Error (MAE). The cross-validation preserved temporal ordering using KFold with shuffle=False.

For XGBoost, the optimization explored: n_estimators (100-800), learning_rate (0.01-0.3, log-scale), max_depth (3-10), min_child_weight (1-10), gamma (1e-8 to 1.0, log-scale), subsample (0.6-1.0), colsample_bytree (0.6-1.0), reg_alpha (1e-8 to 10.0, log-scale), and reg_lambda (1e-8 to 10.0, log-scale).

Step 3: Performance Evaluation and Model Selection

We assessed model performance using five metrics: (1) R-squared measuring variance explained; (2) Mean Absolute Error (MAE) quantifying average prediction error; (3) Root Mean Squared Error (RMSE) penalizing large errors; (4) Mean Absolute Percentage Error (MAPE) providing relative error; (5) Comprehensive Score combining test R-squared (40% weight), test MAE (30%), generalization gap between CV and test MAE (20%), and CV MAE (10%). Scores range from 0 to 100. Results are showed in **SI-Figure 23-27.**

Step 4: XGBoost Algorithm and Optimization

XGBoost Selection: Based on comprehensive evaluation, XGBoost was selected as the optimal model. Key metrics: Cross-Validation MAE 0.6237 (± 0.0090), Test R-squared 0.8541, Test MAE 0.3072, Test RMSE 0.4309, Test MAPE 111.79%, Generalization Gap 0.3164. **Optimized Configuration:** n_estimators 736, learning_rate 0.0513, max_depth 4, min_child_weight 7, gamma 0.000949, subsample 0.763, colsample_bytree 0.942, reg_alpha 0.297, reg_lambda 8.547.

XGBoost employs an additive model aggregating K decision trees. The training objective

minimizes a regularized loss function balancing prediction accuracy and model complexity. At each iteration, XGBoost applies second-order Taylor expansion around previous predictions, exploiting both first-order gradients and second-order gradients (Hessian) for precise optimization. For candidate splits, the algorithm calculates gain measuring reduction in loss, executing splits only when gain exceeds the complexity penalty gamma. This automatically balances model fit and complexity, producing trees that generalize well.

XGBoost's superiority stems from: (1) Gradient boosting framework iteratively correcting residual errors; (2) Built-in L1 and L2 regularization preventing overfitting (training R-squared 0.9912 vs test R-squared 0.8541); (3) Second-order gradient information enabling faster convergence; (4) Industrial maturity with robust implementation.

Step 5: Model Interpretability via SHAP Values

To interpret the XGBoost model, we employed SHAP (SHapley Additive exPlanations), grounded in cooperative game theory. SHAP values decompose predictions into additive feature contributions. For feature i, the SHAP value quantifies its marginal contribution across all possible feature coalitions, weighted to ensure equal treatment of all orderings. The sum of all SHAP values equals the difference between the model's prediction and global average, providing complete attribution.

SHAP values provide a rigorous method to decompose any model's prediction into additive contributions from individual features. For feature i, its SHAP value ϕ_i quantifies the feature's marginal contribution across all possible feature coalitions:

$$\phi_i = \sum_{S \subset N,\ \{i\}} \frac{|S|!(M-|S|-1)!}{M!}[f(S \cup \{i\}) - f(S)] \tag{0.1}$$

where:

- N: complete feature set
- M: total number of features (M=25 in our case)
- S: a subset of features excluding feature i
- $f(S)$: model's prediction using only features in subset S
- $f(S \cup \{i\})$: prediction after adding feature i to subset S

The weighting factor $\frac{|S|!(M-|S|-1)!}{M!}$ ensures equal treatment of all possible orderings in which features could be added to the coalition.

We used the TreeSHAP algorithm, which exploits tree structure to reduce computational complexity from exponential to polynomial time, making SHAP analysis feasible for our 736-tree XGBoost model with 14 features. We also conducted a dependence analysis of all feature variables in **SI-Figure 28-32**.

Step 6: Reproducibility and Data Availability

To ensure full reproducibility: (1) Fixed random seed (seed=42) across all components; (2) Single-threaded execution (n_jobs=1), eliminating parallelization variability; (3) Deterministic Bayesian optimization via Optuna TPESampler with seed=42 and multivariate=True; (4) Complete environment specification (Python 3.12, scikit-learn 1.3.0, XGBoost 2.0.0, Optuna 3.3.0, SHAP 0.42.1); (5) Total optimization runtime: 90.09 minutes for all nine models.

5. Data and Code Availability

We use Sims's (2025) maps to measure the extent of forest loss and its drivers [54] as the baseline because it is one of the most recent forest loss datasets powered by AI models and provides

763 information on drivers at a global scale. The MRIO data we used is REX3, a global MRIO database
764 that divides the global economy into 163 representative sectors and covers 189 countries from 1995
765 to 2022. The analysis code is developed in MATLAB and Python. The geospatial maps are created
766 through Python. The code used in this study will be made available upon request. Data sources of 14
767 driving factors include World Bank National Accounts, OECD Statistics, UN Population Division,
768 World Database on Protected Areas (WDPA), and IMF International Financial Statistics. The yearly
769 country-to-country forest loss footprint flows dataset from 2001 to 2022 is available in the
770 supplementary Excel 1.
771
772

6. Reference

1 Pan, Y., Birdsey, R. A., Phillips, O. L. & Jackson, R. B. The Structure, Distribution, and Biomass of the World's Forests. *Annual Review of Ecology, Evolution, and Systematics* **44**, 593-622 (2013).

2 Jenkins, C. N., Pimm, S. L. & Joppa, L. N. Global patterns of terrestrial vertebrate diversity and conservation. *Proceedings of the National Academy of Sciences* **110**, E2602-E2610 (2013).

3 Georg Kappen, E. K., Torsten Kurth, Johanna Puetz, Andreas Reinhardt, and Juuso Soininen. The Staggering Value of Forests—and How to Save Them. https://www.bcg.com/publications/2020/the-staggering-value-of-forests-and-how-to-save-them (2020).

4 Hickey,V. Forests for healthy people, economies, and ecosystems, https://blogs.worldbank.org/en/voices/forests-healthy-people-economies-and-ecosystems (2023).

5 Michelle Sims, R. S., Maxim Neumann, Anton Raichuk, Drew Purves. New Data Shows What's Driving Forest Loss Around the World, https://www.wri.org/insights/forest-loss-drivers-data-trends (2025).

6 Besnard, S., Heinrich, V. H. A., Carvalhais, N., Ciais, P., Herold, M., Luijkx, I., Peters, W., Requena Suarez, D., Santoro, M. & Yang, H. Global covariation of forest age transitions with the net carbon balance. *Nature Ecology & Evolution* **9**, 1848-1860 (2025).

7 Pan, Y., Birdsey, R. A., Phillips, O. L., Houghton, R. A., Fang, J., Kauppi, P. E., Keith, H., Kurz, W. A., Ito, A., Lewis, S. L., Nabuurs, G.-J., Shvidenko, A., Hashimoto, S., Lerink, B., Schepaschenko, D., Castanho, A. & Murdiyarso, D. The enduring world forest carbon sink. *Nature* **631**, 563-569 (2024).

8 FAO. The State of the World's Forests 2024– Forest-sector innovations towards a more sustainable future. Report No. 978-92-5-138867-9, (Rome, 2024).

9 Marsh, C. J., Turner, E. C., Blonder, B. W., Bongalov, B., Both, S., Cruz, R. S., Elias, D. M. O., Hemprich-Bennett, D., Jotan, P., Kemp, V., Kritzler, U. H., Milne, S., Milodowski, D. T., Mitchell, S. L., Pillco, M. M., Nunes, M. H., Riutta, T., Robinson, S. J. B., Slade, E. M., Bernard, H., Burslem, D. F. R. P., Chung, A. Y. C., Clare, E. L., Coomes, D. A., Davies, Z. G., Edwards, D. P., Johnson, D., Kratina, P., Malhi, Y., Majalap, N., Nilus, R., Ostle, N. J., Rossiter, S. J., Struebig, M. J., Tobias, J. A., Williams, M., Ewers, R. M., Lewis, O. T., Reynolds, G., Teh, Y. A. & Hector, A. Tropical forest clearance impacts biodiversity and function, whereas logging changes structure. *Science* **387**, 171-175 (2025).

10 Xu, X., Fu, D., Su, F., Lyne, V., Yu, H., Tang, J., Hong, X. & Wang, J. Global distribution and decline of mangrove coastal protection extends far beyond area loss. *Nature Communications* **15**, 10267 (2024).

11 Global Warming of 1.5°C. An IPCC Special Report on the impacts of global warming of 1.5°C above pre-industrial levels and related global greenhouse gas emission pathways, in the context of strengthening the global response to the threat of climate change, sustainable development, and efforts to eradicate poverty., (Intergovernmental Panel on Climate Change (IPCC)), https://www.ipcc.ch/sr15/.

12 Villegas, A. World falling far behind deforestation goals with farms and fires driving loss, report says, https://www.reuters.com/sustainability/climate-energy/world-falling-far-behind-

deforestation-goals-with-farms-fires-driving-loss-2025-10-14/ (2025).
13 Ritchie, H. Drivers of Deforestation. Our World in Data, https://ourworldindata.org/drivers-of-deforestation (2021).
14 Gullison, R. E., Frumhoff, P. C., Canadell, J. G., Field, C. B., Nepstad, D. C., Hayhoe, K., Avissar, R., Curran, L. M., Friedlingstein, P., Jones, C. D. & Nobre, C. Tropical Forests and Climate Policy. *Science* **316**, 985-986 (2007).
15 Seymour, F. & Harris, N. L. Reducing tropical deforestation. *Science* **365**, 756-757 (2019).
16 Villoria, N., Garrett, R., Gollnow, F. & Carlson, K. Leakage does not fully offset soy supply-chain efforts to reduce deforestation in Brazil. *Nature Communications* **13**, 5476 (2022).
17 Heilmayr, R., Carlson, K. M. & Benedict, J. J. Deforestation spillovers from oil palm sustainability certification. *Environmental Research Letters* **15** (2020).
18 Carlson, K. M., Heilmayr, R., Gibbs, H. K., Noojipady, P., Burns, D. N., Morton, D. C., Walker, N. F., Paoli, G. D. & Kremen, C. Effect of oil palm sustainability certification on deforestation and fire in Indonesia. *Proceedings of the National Academy of Sciences* **115**, 121-126 (2018).
19 Heilmayr, R. & Lambin, E. F. Impacts of nonstate, market-driven governance on Chilean forests. *Proceedings of the National Academy of Sciences* **113**, 2910-2915 (2016).
20 Lambin, E. F., Gibbs, H. K., Heilmayr, R., Carlson, K. M., Fleck, L. C., Garrett, R. D., le Polain de Waroux, Y., McDermott, C. L., McLaughlin, D., Newton, P., Nolte, C., Pacheco, P., Rausch, L. L., Streck, C., Thorlakson, T. & Walker, N. F. The role of supply-chain initiatives in reducing deforestation. *Nature Climate Change* **8**, 109-116 (2018).
21 Pendrill, F., Persson, U. M., Godar, J. & Kastner, T. Deforestation displaced: trade in forest-risk commodities and the prospects for a global forest transition. *Environmental Research Letters* **14**, 055003 (2019).
22 DeFries, R. S., Rudel, T., Uriarte, M. & Hansen, M. Deforestation driven by urban population growth and agricultural trade in the twenty-first century. *Nature Geoscience* **3**, 178-181 (2010).
23 Heilmayr, R., Echeverría, C. & Lambin, E. F. Impacts of Chilean forest subsidies on forest cover, carbon and biodiversity. *Nature Sustainability* **3**, 701-709 (2020).
24 Curtis, P. G., Slay, C. M., Harris, N. L., Tyukavina, A. & Hansen, M. C. Classifying drivers of global forest loss. *Science* **361**, 1108-1111 (2018).
25 Lapola, D. M., Pinho, P., Barlow, J., Aragão, L. E. O. C., Berenguer, E., Carmenta, R., Liddy, H. M., Seixas, H., Silva, C. V. J., Silva-Junior, C. H. L., Alencar, A. A. C., Anderson, L. O., Armenteras, D., Brovkin, V., Calders, K., Chambers, J., Chini, L., Costa, M. H., Faria, B. L., Fearnside, P. M., Ferreira, J., Gatti, L., Gutierrez-Velez, V. H., Han, Z., Hibbard, K., Koven, C., Lawrence, P., Pongratz, J., Portela, B. T. T., Rounsevell, M., Ruane, A. C., Schaldach, R., da Silva, S. S., von Randow, C. & Walker, W. S. The drivers and impacts of Amazon forest degradation. *Science* **379**, eabp8622(2023)
26 zu Ermgassen, E. K. H. J., Bastos Lima, M. G., Bellfield, H., Dontenville, A., Gardner, T., Godar, J., Heilmayr, R., Indenbaum, R., dos Reis, T. N. P., Ribeiro, V., Abu, I.-o., Szantoi, Z. & Meyfroidt, P. Addressing indirect sourcing in zero deforestation commodity supply chains. *Science Advances* **8**, eabn3132(2022)
27 Runze, L., Jian, Z., Ke, C. & Yang, W. Mechanisms by which environmental regulation and social network embeddedness influence farmers' ecological efficiency. *Scientific Reports* **15**, 35118 (2025).
28 Hoang, N. T. & Kanemoto, K. Mapping the deforestation footprint of nations reveals growing

threat to tropical forests. *Nature Ecology & Evolution* **5**, 845-853 (2021).
29 Hoang, N. T., Taherzadeh, O., Ohashi, H., Yonekura, Y., Nishijima, S., Yamabe, M., Matsui, T., Matsuda, H., Moran, D. & Kanemoto, K. Mapping potential conflicts between global agriculture and terrestrial conservation. *Proceedings of the National Academy of Sciences* **120**, e2208376120 (2023).
30 West, C., Rabeschini, G., Singh, C., Kastner, T., Bastos Lima, M., Dermawan, A., Croft, S. & Persson, U. M. The global deforestation footprint of agriculture and forestry. *Nature Reviews Earth & Environment* **6**, 325-341 (2025).
31 Liang, J., Wang, S., Liu, K., Wang, J., Zhou, J., Liu, X., Feng, K. & Fang, C. Agricultural HANPP Embodied in Consumption: Tracing Pressure on Ecosystems Based on an MRIO Analysis. *Environmental Science & Technology* **57**, 13838-13850 (2023).
32 Yamano, N., Alsamawi, A., Webb, C., Cimper, A., Zürcher, C. & Chiapin Pechansky, R. Development of the OECD Inter Country Input-Output Database 2023. Report No. 18151965, (OECD 2023).
33 Pendrill, F., Gardner, T. A., Meyfroidt, P., Persson, U. M., Adams, J., Azevedo, T., Bastos Lima, M. G., Baumann, M., Curtis, P. G., De Sy, V., Garrett, R., Godar, J., Goldman, E. D., Hansen, M. C., Heilmayr, R., Herold, M., Kuemmerle, T., Lathuillière, M. J., Ribeiro, V., Tyukavina, A., Weisse, M. J. & West, C. Disentangling the numbers behind agriculture-driven tropical deforestation. *Science* **377**, eabm9267(2022)
34 Wang, Y., Hollingsworth, P. M., Zhai, D., West, C. D., Green, J., Chen, H., Hurni, K., Su, Y., Warren-Thomas, E., Xu, J. & Ahrends, A. New high-resolution maps show that rubber causes significant deforestation. *bioRxiv*, 2022.2012.2003.518959 (2022).
35 Wang, Y., Hollingsworth, P. M., Zhai, D., West, C. D., Green, J. M. H., Chen, H., Hurni, K., Su, Y., Warren-Thomas, E., Xu, J. & Ahrends, A. High-resolution maps show that rubber causes substantial deforestation. *Nature* **623**, 340-346 (2023).
36 Taheripour, F., Zhao, X. & Tyner, W. E. The impact of considering land intensification and updated data on biofuels land use change and emissions estimates. *Biotechnology for Biofuels* **10**, 191 (2017).
37 Bjelle, E. L., Többen, J., Stadler, K., Kastner, T., Theurl, M. C., Erb, K.-H., Olsen, K.-S., Wiebe, K. S. & Wood, R. Adding country resolution to EXIOBASE: impacts on land use embodied in trade. *Journal of Economic Structures* **9**, 14 (2020).
38 Zhang, X., Chen, J., Lenzen, M., Brunet-Navarro, P., Dias, A. C., Shao, S., Liu, G., Cao, Z., Hu, W., Lu, F., Yang, H. & Ouyang, Z. Carbon Balance of China-Made Wood Products Assessed Using a Trade-Linked Approach. *Environmental Science & Technology* **59**, 10227-10238 (2025).
39 Zhang, X., Chen, J., Dias, A. C. & Yang, H. Improving Carbon Stock Estimates for In-Use Harvested Wood Products by Linking Production and Consumption—A Global Case Study. *Environmental Science & Technology* **54**, 2565-2574 (2020).
40 Timmer, M. P., Dietzenbacher, E., Los, B., Stehrer, R. & de Vries, G. J. An Illustrated User Guide to the World Input–Output Database: the Case of Global Automotive Production. *Review of International Economics* **23**, 575-605 (2015).
41 Lenzen, M., Kanemoto, K., Moran, D. & Geschke, A. Mapping the Structure of the World Economy. *Environmental Science & Technology* **46**, 8374-8381 (2012).
42 Wang, S., Chen, X., Xie, R., Liu, K., Wang, J., Liu, X., Hubacek, K., Wu, C., Feng, K., Yan,

Y., Liu, Z., Sun, L. & Fang, C. Demand-side insights for steering human appropriation of net primary productivity within planetary boundaries. *One Earth* **7**, 650-662 (2024).

43 Timmer, M. P. L., Bart ; Stehrer, Robert et al. An Anatomy of the Global Trade Slowdown based on the WIOD 2016 Release. https://research.rug.nl/en/publications/an-anatomy-of-the global-trade-slowdown-based-on-the-wiod-2016-re/ (2016).

44 Corong, U. L. C. B. a. E. L. Development of GTAP version 11Land Use and Land Cover database for years 2004, 2007, 2011, 2014 and 2017. https://www.gtap.agecon.purdue.edu/resources/res_display.asp?RecordID=7407 (2025).

45 Pendrill, F., Persson, U. M., Godar, J., Kastner, T., Moran, D., Schmidt, S. & Wood, R. Agricultural and forestry trade drives large share of tropical deforestation emissions. *Global Environmental Change* **56**, 1-10 (2019).

46 Moran, D. & Kanemoto, K. Identifying species threat hotspots from global supply chains. *Nature Ecology & Evolution* **1**, 0023 (2017).

47 Bruckner, M., Wood, R., Moran, D., Kuschnig, N., Wieland, H., Maus, V. & Börner, J. FABIO-The Construction of the Food and Agriculture Biomass Input-Output Model. *Environ Sci Technol* **53**, 11302-11312 (2019)

48 Sims, M. J., Stanimirova, R., Raichuk, A., Neumann, M., Richter, J., Follett, F., MacCarthy, J., Lister, K., Randle, C., Sloat, L., Esipova, E., Jupiter, J., Stanton, C., Morris, D., Melhart Slay, C., Purves, D. & Harris, N. Global drivers of forest loss at 1 km resolution. *Environmental Research Letters* **20**, 074027 (2025).

49 Cazcarro, I., Usubiaga-Liaño, A., Román, M. V., Piñero, P., Dietzenbacher, E., Rueda-Cantuche, J. M. & Arto, I. FIGARO-E3: a high-resolution extended multi-regional input-output database consistent with official statistics. *Scientific Data* **12**, 575 (2025).

50 Stadler, K., Wood, R., Bulavskaya, T., Södersten, C.-J., Simas, M., Schmidt, S., Usubiaga, A., Acosta-Fernández, J., Kuenen, J., Bruckner, M., Giljum, S., Lutter, S., Merciai, S., Schmidt, J. H., Theurl, M. C., Plutzar, C., Kastner, T., Eisenmenger, N., Erb, K.-H., de Koning, A. & Tukker, A. EXIOBASE 3: Developing a Time Series of Detailed Environmentally Extended Multi-Regional Input-Output Tables. *Journal of Industrial Ecology* **22**, 502-515 (2018).

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SIExcel1ForestIndatabase.xlsx
- SupportingInformation.docx
- SIExcel2ForestLossFootprintDataset.xlsx