Berlin, July 2024

Making Deforestation Due Diligence Work in Practice

A Practical Methodology & Implementation Guidance for Financial Institutions

BY CLIMATE & COMPANY, AP2, GLOBAL CANOPY

Lead Authors:

Malte Hessenius (malte@climcom.org)

Marc Bohnet (marc@climcom.org)

Tycho Tax (tycho@climcom.org)

Contributing Authors:

Åsa Mossberg (AP2)

Fiona Pedeboy (Global Canopy)

Helen Bellfield (Global Canopy)

Pei Chi Wong (Global Canopy)

Version 1.0

TABLE OF CONTENTS

TABLE OF CONTENTS

EXECUTIVE SUMMARY

Context & Emerging Frameworks

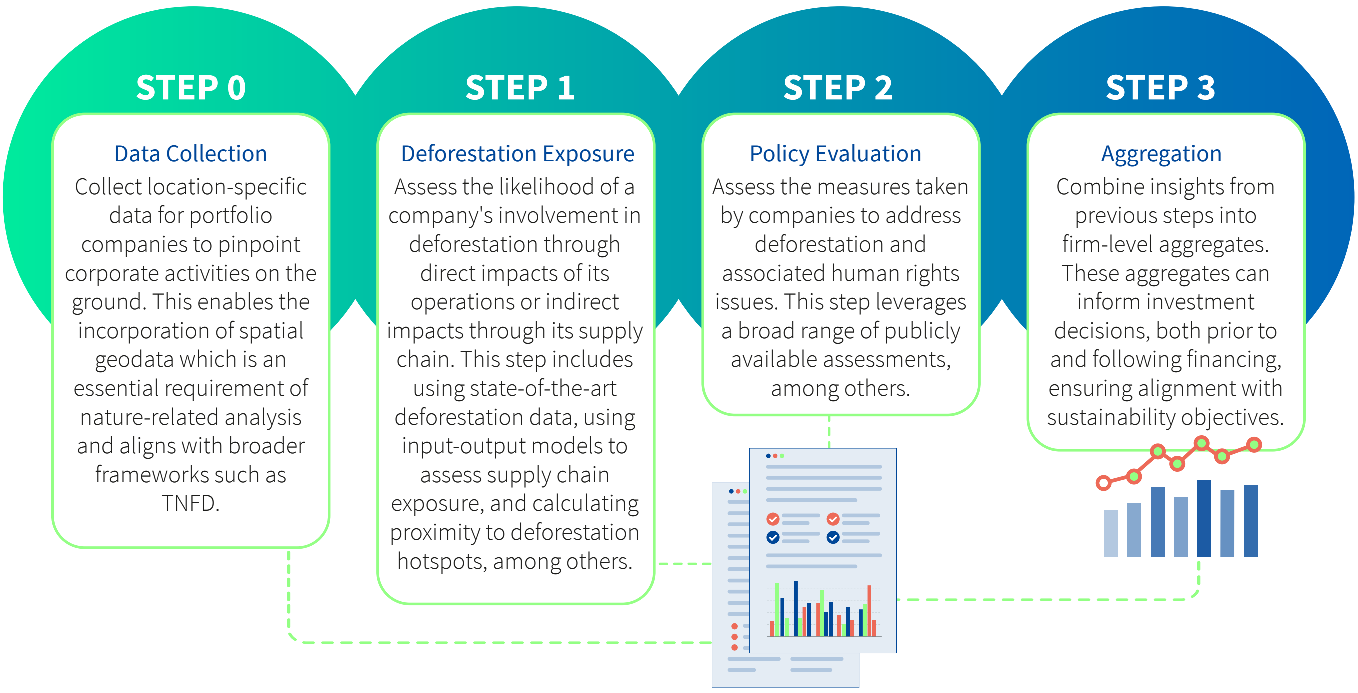
Biodiversity is declining faster than at any time in human history and is predicted to have far-reaching and irreversible consequences for the planet. Deforestation and land conversion are one of the main drivers of biodiversity loss and climate change, and so to address these issues we must make informed choices to reduce deforestation. The financial sector has a key role to play in accelerating the transition towards a deforestation-free economy.

In response to these challenges, global environmental frameworks such as the Global Biodiversity Framework (GBF) are increasingly including the financial sector, creating the need for a robust response. Fortunately, global frameworks are emerging to assess and integrate nature-related issues into financial decision-making, with the Taskforce on Nature-related Financial Disclosure (TNFD) emerging as the global reference framework. New regulations in the European Union, for example, also require the financial sector to conduct a nature-related materiality assessment. All of this requires a robust and practical approach to identifying, preventing and mitigating nature-related risks and impacts, which, despite methodological developments, is still a more complex dimension than climate due to location-specific data requirements.

A Practical Methodology & Implementation Guidance

To address these needs, we have developed a practical methodology tailored for financial institutions conducting large-scale assessments across potentially thousands of portfolio companies. Focusing on deforestation—a critical environmental issue but also a tangible use case for financial institutions to start addressing nature-related aspects —our modular methodology aligns well with broader nature-related frameworks. It has been tested using the MSCI ACWI as a global benchmark and by Andra AP-fonden (AP2) for their listed equities portfolio. Building on Deforestation-free Due Diligence Guidance published by Global Canopy, Neural Alpha, and the Stockholm Environment Institute (2023), this report is complemented by online resources to facilitate implementation and an open-source Excel file with deforestation indicators for the MSCI ACWI.

Our primary objective is to equip financial institutions with actionable steps to make deforestation-free commitments less resource-intensive. By synthesising multiple data sources and leveraging existing guidelines, this report delivers a systematic approach covering:

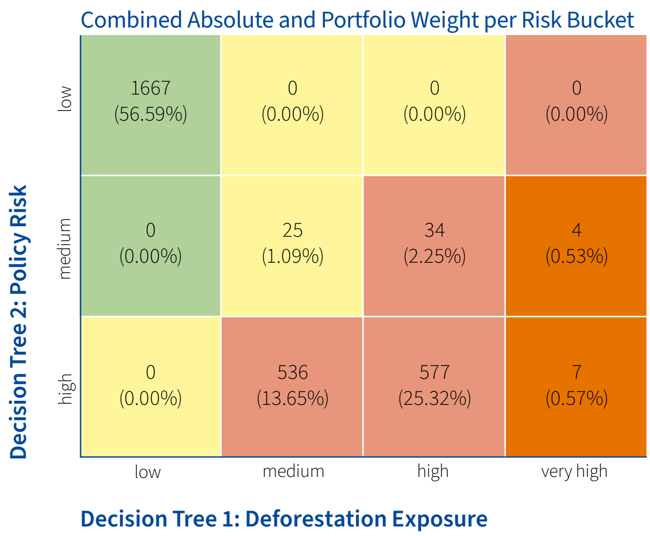


Practical Implications

While one section provides AP2's perspective, and how AP2 has applied this methodology to its listed equity portfolio, we exemplify the analysis for the MSCI ACWI as a global benchmark for equity portfolios. This provides a deep dive into the practical challenges and solutions in managing deforestation risks at scale.

One option to aggregate the results is to overlay the deforestation exposure analysis (split in low to very high risk buckets) with portfolio companies' policy risks (split into low-high buckets), resulting in the following matrix. Practitioners could prioritise the companies in the bottom-right corner to validate results via further scrutiny. These companies are characterised by high deforestation exposure and an insufficient set of policies to mitigate the risks. After validation, these companies could become part of meaningful post-financing decisions such as engagement to maximise investor contribution potential. This analysis can be readily used to report on portfolio metrics that align closely with the TNFD Core Sector Metrics for financial institutions (see **Figure 14**).

Resulting buckets for the MSCI ACWI



Another option is to derive a numerical score that takes full advantage of the depth of the input data. As some companies are scored based on hundreds of data points that provide rich insights into their activities around the world, the score is very heterogeneous and provides an excellent starting point for prioritising portfolio companies or incorporating it into metric-based approaches. (see **Figure 13**)

Complementary Online Resources

This report is also complemented by an online code repository hosted on GitHub (link). The publicly available code repository is based entirely on publicly available data, which can be supplemented by the user's own (proprietary) data if desired. The repository also contains an open source Excel file illustrating the analysis for the MSCI ACWI universe.

INTRODUCTION

Deforestation and growing pressure to act. In recent years, the issue of deforestation has been catapulted to the forefront of global environmental concerns, commanding the attention not only of policymakers and environmentalists but also of financial markets. Debates surrounding the introduction of regulations and frameworks have highlighted the role financial institutions play in either exacerbating or mitigating deforestation. Exemplifying this growing momentum are initiatives such as the European Union's Deforestation Regulation (EUDR) and its review mechanisms for the involvement of financial institutions, the EU's Supply Chain Law, also known as the Corporate Sustainability Due Diligence Directive (CSDDD), the Taskforce on Nature-related Financial Disclosures (TNFD), and the Global Biodiversity Framework (GBF). In the meantime, a growing number of financial institutions have taken a proactive stance to limit deforestation[1].

Why is there a need for another report? In this evolving landscape, key datasets, tools, and guidelines have emerged to help practitioners act on their 'deforestation-free' commitments. Global Canopy has been at the forefront of this development, collecting key datasets (e.g. Forest 500 or the Deforestation Action Tracker) and publishing a how-to guide (in partnership with Neural Alpha and the Stockholm Environmental Institute) that sets out a recommended due diligence approach for financial institutions to identify, prevent and mitigate deforestation-related risks and impacts (Global Canopy et al, 2023) that sets out a recommended due diligence approach for financial institutions to identify, prevent and mitigate deforestation-related risks and impacts. The latter has been a particularly helpful step towards breaking down deforestation-related assessments into concrete, actionable steps. Still, executing the guidance for global portfolios remains time consuming, especially the examination of location-specific data as well as the indirect deforestation-exposure through the value chain of portfolio companies. Furthermore, most tools, while reasonably prioritising high-impact companies, cover only fractions of global equity portfolios. Finally, while the how-to-guide provides detailed guidance to financial institutions, it has been written to accommodate a broad range of use cases and is agnostic to implementation methods in the context of any specific financial institution. It is therefore useful to provide an in-depth real-life example of how it has been applied in practice.

In this context, Climate & Company has partnered with the Swedish pension fund Andra AP-fonden (AP2), a signatory to the Financial Sector Commitment Letter on Eliminating Commodity-driven Deforestation[2], to develop a practical & workable methodology for their equity portfolio which combines existing data sources and systematises the assessment as much as possible. We implement and document each of the steps, so that proposed solutions to technical challenges can be readily adopted by other financial institutions and enable the financial sector as a whole to move forward on deforestation due diligence. This collaboration is actively supported by Global Canopy, through additional expertise and resources. By making our findings complemented by an online repository (see Annexes), we hope to minimise time and staff resources needed to implement deforestation-free commitments (**Figure 1**).

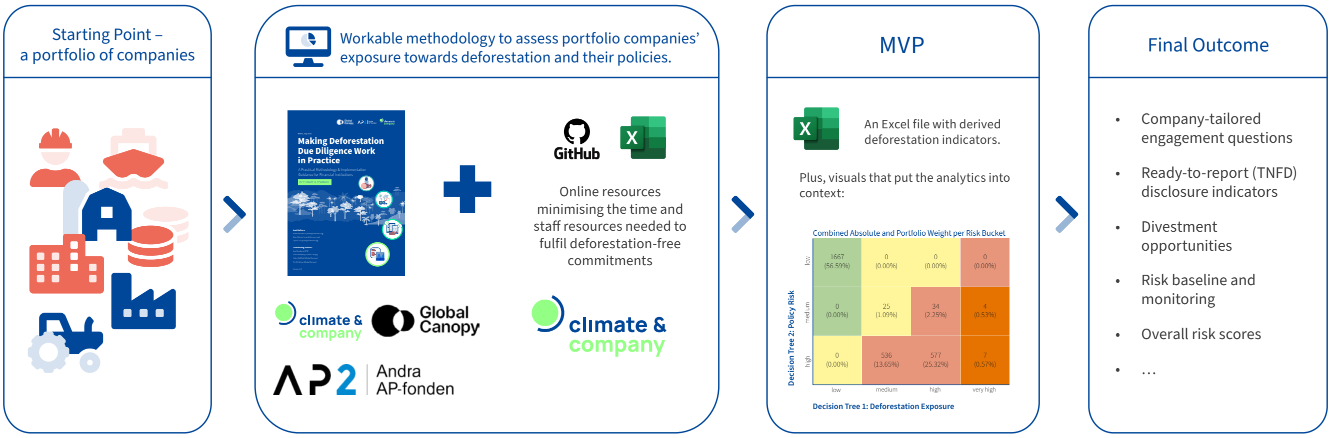
Target audience and asset classes. This guidance, and the associated online resources are designed to be relevant for a broad spectrum of financial institutions. This includes those that have pledged commitment to tackling deforestation and are seeking to enhance their due diligence frameworks, to those yet to prioritise deforestation but are interested in initiating high-impact, minimal-effort measures. It also serves as an excellent deep dive for financial institutions implementing TNFD, since there are large synergies with the TNFD's voluntary assessment approach: the LEAP approach. In its current form, this guidance is particularly suited to portfolios of corporates, in particular global equity, and corporate bond portfolios – though there are steps that can be selectively applied to those covering small and medium-sized enterprises (SMEs). However, as it stands, it is less applicable to investments in non-corporates such as sovereign or sub-sovereign bonds or cash holdings where the "location" and "sector" is less clear.

Synergies with TNFD. A lot of the infrastructure needed to assess deforestation-exposure can likely be reused in other contexts. We therefore highlight synergies with the evolving TNFD framework throughout the report. Drawing parallels to the TCFD, we anticipate that the TNFD will soon gain regulatory endorsement, thereby reshaping mandatory reporting landscapes. Regardless of alignment with the TNFD[3], the

1 Such as the Finance Sector Deforestation Action (FSDA) initiative or the Investors Policy Dialogue on Deforestation (IPDD).

2 See link

Figure 1 - Overview Deliverables



integration of location-specific company and geodata, as well as a systemic look at the value chains of global equities, will be important know-how to navigate the growing challenges of nature-related risks and impacts.

Global Canopy et al.'s approach summarised

Global Canopy et al's (2023) Due Diligence Guidance is organised into two decision trees. Decision Tree 1 (DT1) assesses the exposure of companies to deforestation. It begins with relatively straightforward steps, such as determining whether a company operates in a high-impact sector and checking for known incidents of or involvement in deforestation. As the assessment progresses, it becomes more resource-intensive and detailed, taking into account the regions from which a company sources and whether it operates sites in known deforestation hotspots. Decision Tree 2 (DT2) evaluates the policy actions taken by companies, categorising them based on the ambition level of their policies on deforestation and human rights. Companies are flagged as high policy risk if they lack these policies, and as low policy risk if they possess strong deforestation and human rights policies. Essentially, both decision trees start with relatively straightforward steps that already allow the categorisation of the risk profile of companies, leaving the more complex data gathering for those identified as high-risk.

Our take summarised

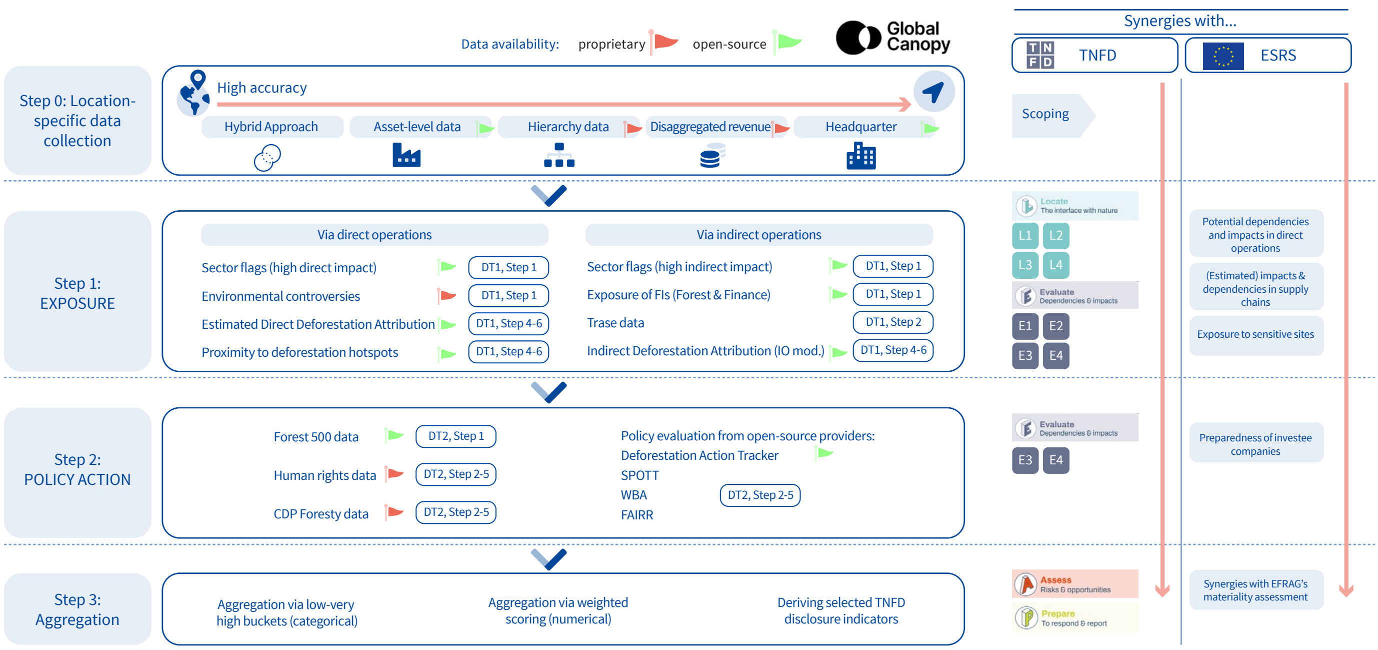
Figure 2 presents a detailed, step-by-step overview of our methodology, linking it to the Global Canopy et al. (2023) guidance and the TNFD's LEAP approach. Our approach follows a similar structure but systematises data gathering and assessment using statistical programming software. This approach facilitates large-scale assessments by enabling assessments regardless of the number of companies involved. Additionally, our methodology introduces an initial data gathering step, termed Step 0, which focuses on using data sources to gather location-specific information on portfolio companies, thereby pinpointing their on-the-ground activities. The methodology then follows Step 1 through to 3, as follows:

- **Step 1 (DT1, exposure evaluation)** clearly distinguishes between direct impacts from the companies' own operations and indirect impacts from their supply chains, as both perspectives come with fundamentally different data needs. This distinction is particularly important for global equities with physical operations primarily in the developed world in order not to severely underestimate their deforestation exposure.
- **Step 2 (DT2, evaluation of policy action)** collects policy data collected from core datasets such as CDP or Forest 500 but also collects information from a whole range of open-source evaluations (such as SPOTT, or the World Benchmarking Alliance).
- **Step 3** elaborates on various ways how the results can be aggregated to derive TNFD-aligned disclosure indicators, and to inform decision making.

[3] Or, for the EU, with EFRAG's materiality guidance. [...]

The remainder of this report demonstrates the application of the methodology using the MSCI ACWI universe as a representative proxy for global equity portfolios. It also includes a section where AP2 applies the methodology to its equity portfolio, providing practical insights and outlining next steps. The last section concludes and provides recommendations for financial institutions, data providers and policy makers.

Figure 2 - Our approach in a nutshell



STEP 0 – COLLECTING LOCATION-SPECIFIC COMPANY DATA

The need to collect location-specific data. Nature-related aspects of issues like deforestation are distinct from purely climate-related matters primarily due to their location-specificity. For example, carbon emissions have a consistent effect on global warming, regardless of where they originate. In contrast, the negative impacts of one hectare of agricultural production can vary greatly depending on its location and whether its impacts natural forest and other natural ecosystems. While forest and ecosystem restoration can help, it cannot fully replicate the biodiversity and ecosystems lost through deforestation, highlighting the critical need for conservation efforts. This underlines the importance of collecting location-specific company data. This data substantially improves the accuracy of the assessment as it links portfolio companies with biodiversity-sensitive areas or high-risk production regions (moving beyond a rather generic sector-country screening). Second, it aligns well with the TNFD framework or even mandated materiality assessments (see ESRS).

Location-specific data: challenges and state of play. TNFD and WWF both mention access to location-specific asset data as one of the key data gaps for green finance (Christiaen, 2023). Moreover, Share Action's assessment of large European banks concludes that it is important to collect portfolio companies' location data (Sood et al., 2022). A World Bank-WWF report has identified the "lack of reliable asset level data at required granularity and regularity" as one of the key data barriers (World Bank & WWF, 2020). "Traditional" asset-level data providers only cover selected economic industries and are rarely intended for nature-related risk analysis (Weber et al., 2017). However, recent efforts have evolved to aggregate and standardise different asset-level datasets. This has resulted in more comprehensive solutions, such as GRESB Asset Impact, making it more useful for practitioners. Also, other larger providers such as Moody's ESG compiled location information internally to derive data products (Christiaen, 2023). Exciting developments are also taking place in the open-source field, ranging from the Global Energy Monitor (covering a range of energy sectors) or the Spatial Finance Initiative (heavy industry and other non-energy related sectors). Additionally, the non-profit initiative Climate Trace has recently received a lot of attention, leveraging satellite imagery to create a global GHG emissions inventory. To conclude, this field is experiencing significant activity and advancements. However, workarounds are still required for large-scale implementation due to limited coverage of industries.

Synergies: This data collection is not only relevant to deforestation but also lays the groundwork for assessing other nature-related risks and human rights issues, including but not limited to the rights of Indigenous Peoples and Local Communities, feeding into the TNFD LEAP approach, or enabling double materiality assessments in the context of the European Sustainability Reporting Standard (ESRS). It also prepares the ground for leveraging information from the emerging field of earth observation data. It is therefore not a stand-alone step but the start of tackling nature-related issues more broadly.

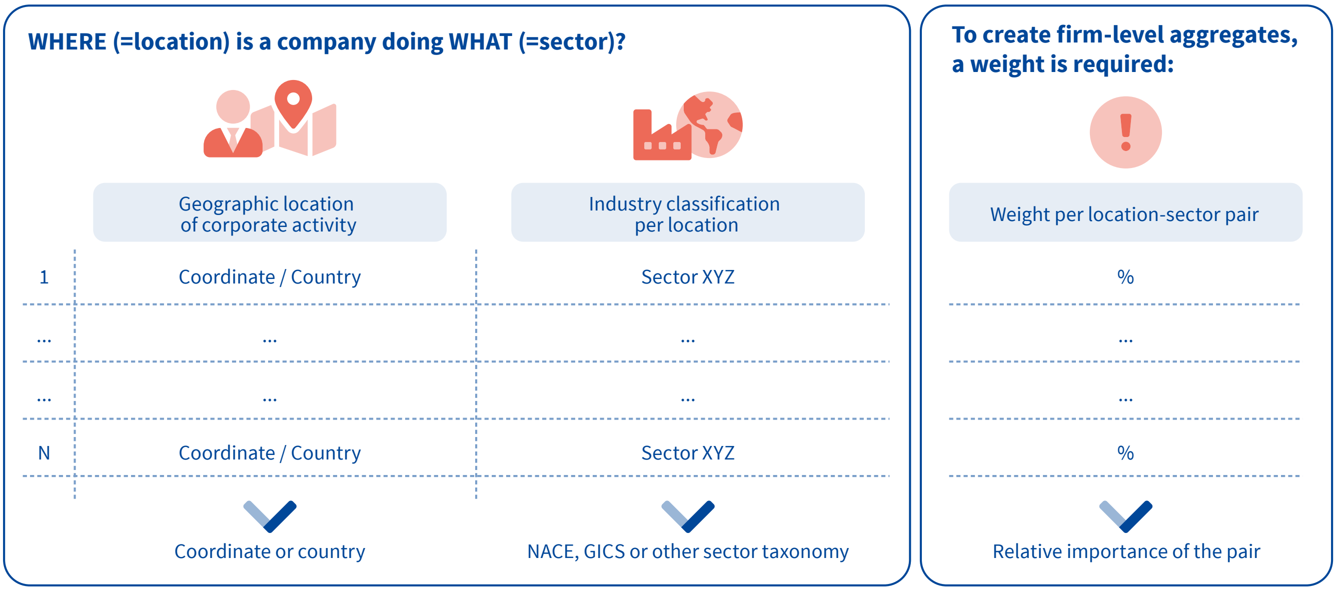
What we do in this report

Step 0 builds on a methodology outlined by the WWF and its Risk Filter Suite[4] - and more information on this methodology is available in the existing report. We do, however, provide resources to facilitate the collection and cleaning of such data in practice (see Annex and Online Repository). The data we use to assess an individual company is a list of "location-sector pairs"[5] (see **Figure 3**). This is crucial, as not only does the location of the company matter but also the type of activity

[4] See WWF and Climate & Company (2023). Tackling Biodiversity Risk – A biodiversity risk assessment guide for companies and financial institutions. (link), or WWF Biodiversity Risk Filter Methodology Documentation (link). The first report provides a high-level summary incl. a case study for a subsample of the MSCI ACWI, whereas the second report contains detailed descriptions on the data collection process for financial institutions (see Guidance A).

[5] Note for the sake of accuracy: different elements of the list of location-sector pairs are used for different purposes. The sectoral distribution is used to derive the share of business activity exposed to high-impact sector, for example. An aggregated version of the list at country-sector level is used for input-output modelling. For an overlay with spatial geodata / deforestation hotspots, we only use the coordinates.

Figure 3 - Location-specific data per company



that is conducted on the ground (e.g., cattle farming would be more relevant than professional services). Together, the two provide a reasonable view of where a company is doing what, which is the foundation for our overall assessment. The location-sector pairs are complemented by an estimate of the relative importance of each site which is required to aggregate the results to the corporate level.

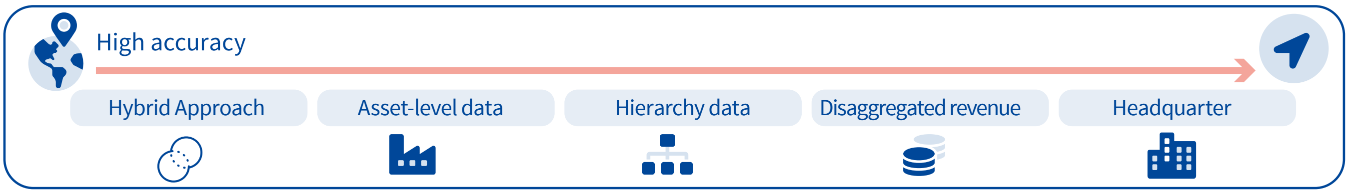
Figure 4 outlines the four data sources that can be used to establish a location-specific data foundation for portfolio companies. Below we describe the pros and cons of each. Note that this data collection approach could also largely be skipped if commercial asset-level data solutions are available in-house[6] For more detailed explanations and specific guidance on how to implement the data collection process, we refer to **Technical Explanation: Collection location-specific company data.** We also provide helpful resources on compiling opensource asset-level data, see **Online Resources: Open-Source Asset-level compilation.**

Asset-level data

Asset-level data refers to specific datasets on physical assets owned by companies including the geographical coordinates of each site. These physical assets could refer to power plants (energy industry), production plants (manufacturing), extractives (mining), farms, and slaughterhouses, among others. These datasets are often complemented by information on production capacity, fuel type or the age of the facility. Primarily, this type of data concentrates on location directly involved in production, such as power plants or extractive industries. Such locations are highly significant for conducting nature-related analyses, as they tend to impact the environment more directly than non-industrial sites like offices.

Historically, such data has been predominantly supplied by commercial providers. However, recent developments in the open-source domain are transforming accessibility. For example, Climate Trace uses satellite data to identify millions of

Figure 4 - Data sources[7]



[6] Such as Asset Resolution (link)

[7] Adapted from WWF (2023). Biodiversity Risk Filter Methodology Documentation, (link)

assets across the globe. Other notable open-source initiatives include the Spatial Finance Initiative, which focuses on sectors like beef abattoirs, petrochemicals, cement, pulp and paper, iron and steel, and waste management. Similarly, the Global Energy Monitor provides extensive coverage of energy-related industries. High quality, open access, satellite data, in combination with advances in image recognition techniques, will likely increase the automated classification of assets. Despite the limited coverage, currently, we estimate that (open source) asset level data will increase in importance in the coming years.

Data on corporate structure and subsidiaries

Accurately assigning subsidiaries to parent companies and navigating through the complexities of company trees is a well-recognised challenge (Worldbank and WWF, 2020). We recommend using corporate hierarchy data, which connects the ultimate parent company to its subsidiaries including information on their location and industry classification. By utilising this data, along with specific data processing steps outlined in the Annex, it is possible to compile a comprehensive list of location-sector pairs for each parent company. This enhances our understanding of the geographical and sectoral distribution of corporate activities, and also allows us to establish corporate responsibility back to the parent company for compliance with investors' expectations.

The primary benefits of this approach include its structured format and the extensive coverage of companies. However, a notable limitation is that the data focuses on corporate structures, so production facilities are only included if they are registered as distinct legal entities.

Disaggregated revenue

The revenue distribution of companies provides an additional, complementary perspective. If reported, disaggregated revenue data breaks down a company's revenue by geographical region (for example, Company A generates 20% of its revenue in Country X) and sector (such as Company A obtaining 10% of its revenue from Sector Y). This segmentation offers only a rough estimate, with the highest level of spatial detail being at the country level. This granularity may serve as an accurate representation for smaller countries like Luxembourg but falls short for larger nations like Brazil. Primarily available for publicly traded companies, these detailed revenue datasets can be sourced from various commercial data providers. This type of data has been used in several biodiversity-related risk assessments. For instance, tools designed for measuring biodiversity impact, like the Corporate Biodiversity Footprint, and studies, such as the one conducted by the Banque de France, have used revenue splits by country and industry sector as the starting point to link companies to models such as GLOBIO or EXIOBASE (WWF 2021, Banque de France 2020). This adds another layer to the analysis, focusing on the (downstream) revenue distribution of larger multinationals in particular.

Headquarter data (location + primary sector)

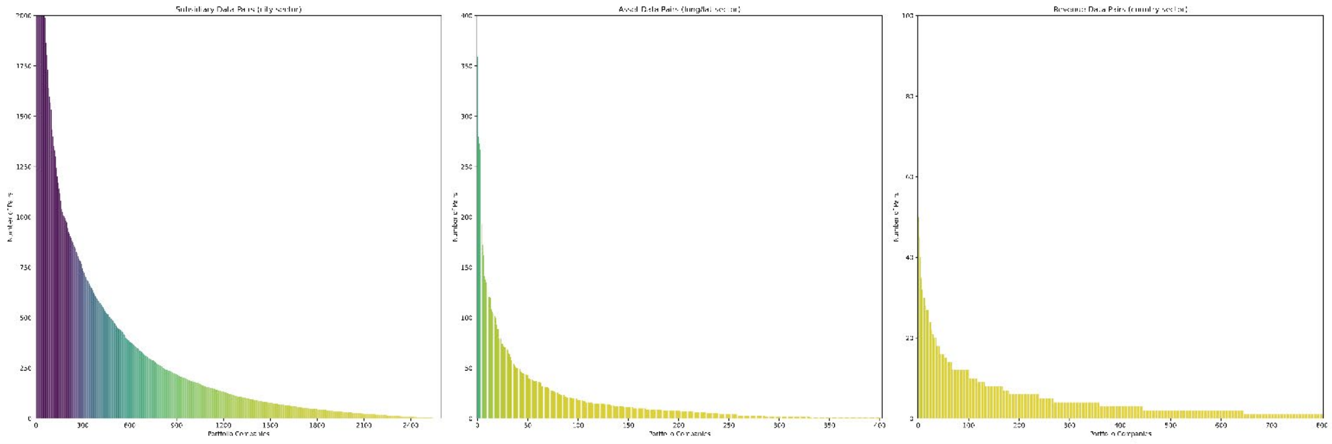
Using the geographic location of the company's registered headquarter as well as the company's primary sector classification can serve as a starting point if no other information is available. This datapoint is easily accessible in countless commercial datasets, as part of fundamental company information and is accessible for most listed and many non-listed companies worldwide.

However, the usefulness of this information heavily depends on the assumption that a company's production activities are closely tied to its headquarters – which often is not the case for larger downstream companies. Therefore, headquarters information should be used with caution, recognising its significant limitations. For smaller companies operating from a single location, this data can sometimes serve as a reasonable proxy.

Combining the different datasets (Hybrid approach)

The steps outlined complement each other effectively, each addressing different aspects of our data needs. Asset-level data provides detailed, location-specific insights into a company's direct business activities. However, its limited coverage means that including location and sector-specific information of subsidiaries can enhance our understanding. By integrating corporate hierarchy data, we can construct a more precise depiction of a company's business activities. For multinational companies that own foreign subsidiaries involved in upstream supply chain activities, the location and industry affiliation of these subsidiaries offer valuable insights into the sectors and locations of their upstream operations. Furthermore, disaggregated revenue information helps researchers understand the downstream activities of companies more thoroughly.

Figure 5 -Data availability per source

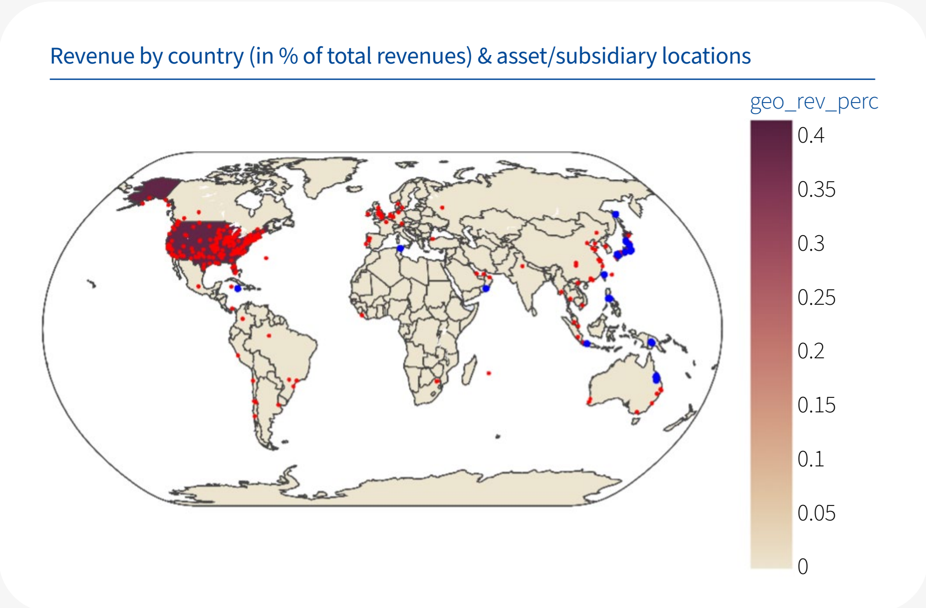


Coverage. Overall, our comprehensive approach has enabled us to collect data on more than 8,000 assets linked to 407 companies (see **Figure 5**, left-hand side). Our approach relies entirely on open-source data, enhancing the reproducibility of our analysis for all users. We have also connected over 800,000 subsidiaries to 2,685 companies (middle) and attributed approximately 5,000 disaggregated country-sector revenue pairs to 904 companies (right). This extensive data collection provides a robust foundation for analysing corporate structures and operations (see Box below for more insights).

BOX: Outcome of Step 0 (Data foundation)

For the exercise in this report, we collected open-source asset-level data (from Climate Trace, Spatial Finance Initiative, and Global Energy Monitor), corporate hierarchy data (ORBIS), disaggregated revenue data (Refinitiv), and headquarters information. **Figure 6** visualises the data collection process for one sample company, where the blue dots are derived from asset level datasets and the red dots from corporate hierarchy data. The different shading per country depicts the revenue distribution.

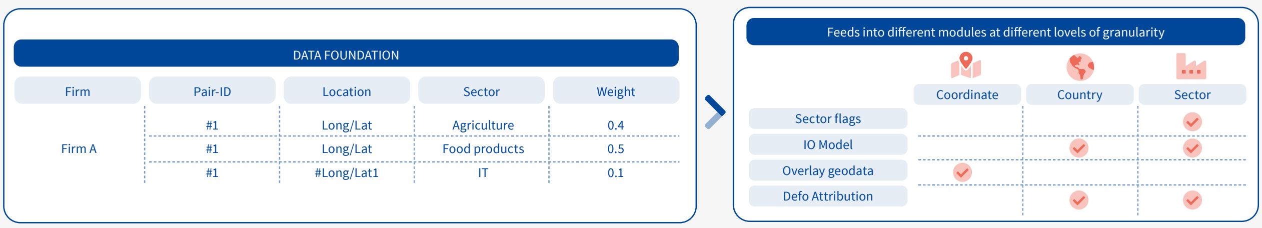
Figure 6 - Sample Company



We conducted the most thorough analysis available by prioritising data sources based on their accuracy (as shown in Figure 4) and aggregating all available data for each company. We enable users to tailor the compilation of these data sources, with the default setting averaging the importance of each country-sector pair across different datasets. For example, consider company A for which only corporate structure and disaggregated revenue data are available. If 10% of all subsidiaries operate in the mining sector in Peru, but only 2% of revenues are generated from this sector, then we would estimate that 6% of company A's operations are dependent on mining in Peru. This method allows us to later assign country-sector specific deforestation risk scores to companies in a much more detailed and precise manner. For more technical details, please refer to the Annex.

From a data processing point of view, we retrieve a list of location-sector per company, which is fed into the different modules, with slight deviations. To derive sector flags, only the sectoral distribution is used, for example (result: X% of the company's revenue is linked to high-impact sectors). For the supply chain analysis, since IO models operate at a country-sector logic, the different pairs are aggregated from location-sector level to country-sector level. For the overlay with deforestation hotspots, we only extract the coordinates, etc.

Figure 7 - Draft illustration data foundation



STEP 1 – DEFORESTATION EXPOSURE

This step operationalises the methodology outlined in Decision Tree 1 of Global Canopy et al. (2023), which provides a structured framework for assessing a company's exposure to deforestation. We differentiate between the two by recognising the distinct data requirements for assessing impacts from direct operations versus those from the supply chain.

1.1. Indicators of Direct Deforestation Exposure

Direct Sectoral Flags [GC Decision Tree1, Step 1]

We use sectoral flags collected from literature and policy reports. These direct flags serve as negative proxies, highlighting industries likely to be exposed to deforestation through their direct operations. Although sectoral flags offer a straightforward method for identifying potential exposure, the assumption of industry-wide homogeneity underpinning these flags is debatable. Despite this, they provide an initial insight into an industry's deforestation exposure, making them a valuable starting point. Furthermore, we use open-access knowledge and research findings to inform our analysis while relying heavily on previous findings by Global Canopy (2021) (European Commission, 2022). All direct and indirect flags can be found in Annex Step 1. Applied to disaggregated firm-level data, the direct sector flags can be used to determine the share of economic activity per company that is potentially exposed to deforestation.

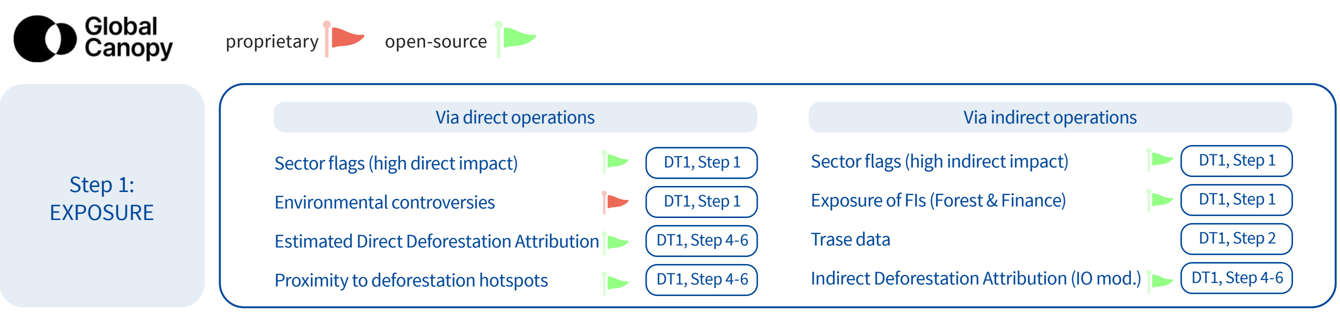
Overlay With Geodata and Direct Exposure [GC Decision Tree 1, Step 4-6]

This step is crucial for the analysis of nature-related risks. While a company's sector classification may provide some initial insights, integrating the geographical location of its activities offers essential context. For instance, a company might be flagged based on its sector alone. However, when this information is combined with geospatial data, it may become clear that the company's physical locations are located far from deforestation hotspots or intact primary forests, significantly altering the risk assessment. To perform the overlay with geospatial data we use the coordinates of the physical assets as well as the location of incorporation of the subsidiaries (see Step 0) before the aggregation step to a list of (sector, region) pairs.

In our analysis we overlaid corporate locations with the following spatial geodata:

- **Direct deforestation exposure data:** Ideally, one would be able to attribute deforestation directly to the activities of each specific company. Aside from data limitations, such as the limited availability of asset-ownership data, such an attribution would depend on an accurate assessment of the sourcing region per asset. Given these considerations, we currently make use of a dataset constructed by Pendrill et al. (2019) and updated by Singh and Persson (2024) which assigns hectares of deforestation to specific (sector,

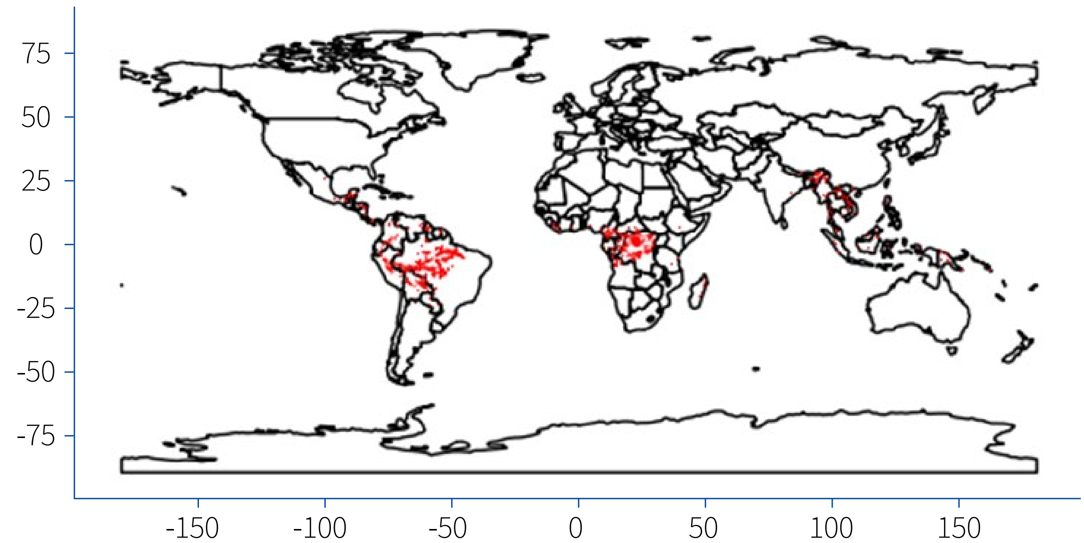
Figure 8 - Step 1, Deep Dive



region) pairs. The original Pendrill dataset (v 1.1) was constructed using spatial datasets on tree cover loss and a land balance model[8] to attribute detected forest loss to agricultural and forest commodities. The recently updated data (v2) is generated by a new model called Deforestation Driver and Carbon Emission (DeDuCE) which is a) more precise (because attribution is to a larger extent based on spatial data rather than on national-level statistics), b) more transparent (because the authors provide a data quality score), c) has global coverage (rather than focussing only on the tropics), and d) has better temporal scope (extended from 2018 to 2022). Further technical details as well as a .csv file with the deforestation allocation can be found in **Annex 1**. Combining this dataset with location specific company data allows one to harness industry-wide averages while still getting a unique deforestation exposure based on the list of (sector, region) pairs, and their weights, for each portfolio company (see Step 0). The deforestation allocation data used in this step is neither perfect as it is partly based on a model of land use change, nor complete given that it only allocates deforestation to tropical agricultural production, cattle meat, forestry products, and peatland drainage and not to certain other high-risk sectors such as direct logging, or mining. The latter is of particular interest in order not to underestimate deforestation exposure for certain companies. We therefore advise to also include both direct and indirect sectoral flags until data and methodological improvements allow for an even more comprehensive estimation of company-specific deforestation exposure in the future[9].

- **Deforestation hot spots [GC Decision Tree 1, Step 5 & 6]:** For a more refined analysis, we used data from Global Forest Watch (2023, GFW) on emerging deforestation hot spots which identifies clusters of primary forest loss. The geodata has been pre-processed, which makes it easy to work with. Deforestation hotspots are classified by GFW as diminishing, intensifying, new, persistent, or sporadic. We apply a different impact weight depending on the type of hotspot (i.e. an intensifying hotspot is assigned a worse score than one which is diminishing) and use a distance threshold to assign physical assets and subsidiaries as being in near proximity to deforestation hot spots or not. The resulting deforestation hot spots are presented in **Figure 9** below. For future improvements, other datasets such as Tree Cover Loss by Dominant Driver[10] could be added as well.

Figure 9 - Deforestation Hotspots (Source: Global Forest Watch)



[8] Firstly, that if cropland expands, it first does so into pastures and then into forests, and secondly that if pastures and forest plantation areas expand, they replace forest.

[9] See for instance the project "ForestNet: Classifying Drivers of Deforestation in Indonesia using Deep Learning on Satellite Imagery" by the Standford ML Group (link).

[10] https://hub.arcgis.com/documents/gfw::tree-cover-loss-by-dominant-driver-2022/about

Environmental Controversies [GC Decision Tree 1, Step 1]

Controversy screenings are another data source to consider as these data products leverage news from third parties to report on corporate involvements in ESG-related controversies. To assess the potential for a company to become entangled in (future) environmental controversies, data can be leveraged from third party ESG data providers such as Bloomberg, MSCI, RepRisk or Refinitiv. Typically, they provide data points on recent controversies (within the last fiscal year) concerning environmental or human rights issues connected to the company's activities. While manual screening of controversies is an option that can provide in-depth and focused insight into deforestation-related issues, we recommend that financial institutions first exploit all targeted information available from the previously mentioned data providers concerning deforestation and human rights. Subsequently, a manual review can be employed to refine the quality of controversy data for a select group of companies. In a not-too-distant future, automated natural language processing (NLP) approaches could provide a scalable alternative to both data from proprietary data providers as well as manual screening. In our MSCI ACWI pilot, we leveraged data from Refinitiv (see **Annex**).

1.2. Indicators of Indirect Deforestation Exposure

While certain economic activities by companies are associated with deforestation directly, many more contribute to deforestation indirectly. It is useful to distinguish between two different types of indirect exposure. The first is indirect exposure through the value chain of a company. The second type of indirect exposure would be through the financing (loans, debt/equity holdings, etc) by financial institutions that allow for corporate activities to take place that either directly or indirectly (through the supply chain) contribute to deforestation. In this chapter we will cover the two different types of indirect exposure sequentially and, while we discuss both types, it is important to keep the distinction between them in mind.

Indirect Sectoral Flags [GC Decision Tree 1, Step 1]

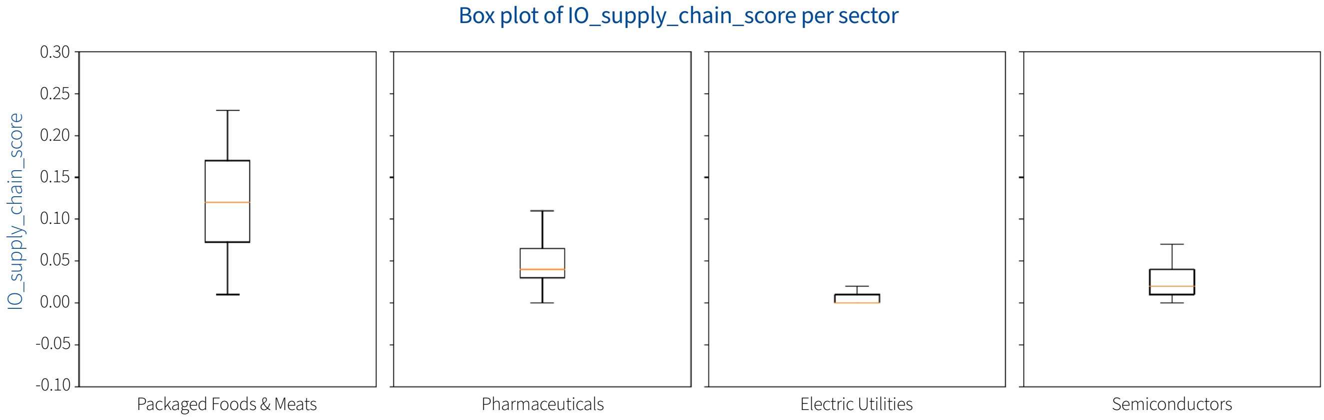
Like direct flags, we use indirect sectoral flags to highlight industries likely to be exposed to deforestation through their supply chain. Given the fact that modern, globalised supply chains are highly complex and interconnected, it should be obvious that indirect sectoral flags are imprecise and should not be used in isolation. In the Finance Sector Roadmap published by Global Canopy (2021), with which the due diligence guidance is aligned, certain financial sector classes (Regional Banks, and Diversified Banks) were flagged as high deforestation risk for the purposes of sector screening. At first glance this is understandable as it is indeed the case that certain financial institutions, through their loans and equity and bond holdings, contribute heavily to deforestation-risk activities. However, given the heterogeneity within the sector, without additional exposure data feeding into the risk assessment, this leads to many false positives, which essentially dilutes the signal of this variable. Fortunately, in the case of the financial sector, the Forest & Finance dataset contains comprehensive exposure data for individual financial institutions. We have therefore decided not to assign a negative indirect sectoral flag for all financial sector codes, and, instead, add a company-specific score that deals with financial institutions separately. All direct and indirect flags can be found in **Annex Step 1**.

Indirect Deforestation Exposure via Input-Output Modelling [GC Decision Tree 1, Step 4-6]

Many goods and services do not directly cause deforestation, but they contribute to it indirectly through their supply chains. Pendrill et al. (2019) estimate that 29–39% of deforestation-related CO_2 emissions are driven by international trade, highlighting the significant (indirect) impact that trade has on deforestation[11]. In theory, to accurately assess these supply chain impacts and link them to portfolio companies, one would take each portfolio company, establish its supply chain relationships, and analyse the activities of each of its suppliers.

[11] It should be noted that trade in agricultural commodities is not the same as the processing of such commodities in downstream products or services, but the figure is nonetheless indicative of the fact that to only assess direct deforestation attribution would be severely underestimating the true effect that downstream companies have in driving deforestation.

Figure 10 - Box plots for four GICS sectors showing how the averages and skewness of the IO score differs significantly.



Unfortunately, in our experience, supply chain data lacks the quality and coverage required for this purpose. The reason is that the disclosure of such relationships is mostly voluntary, and due to its potential sensitivity, not many companies decide to publish them. To still get an estimate of the indirect exposure we rely on EXIOBASE3: a state-of-the-art environmentally extended multi-regional input-output model (MRIO). This model allows us to 'propagate' the direct deforestation allocation data across borders, throughout the global economy. Technical details on how MRIOs work can be found in the original EXIOBASE3 publication[12]. Implementation guidance on how to use deforestation data in combination with EXIOBASE3 can be found in **Annex 1**.

For the purposes of this report, it suffices to clarify that the MRIO provides an estimate of how each (sector, region) pair is connected to all other (sector, region) pairs in the economy. When combined with deforestation exposure data, this model gives us an estimate of the average indirect exposure to deforestation for each downstream (sector, region) pair, through its relationships with upstream (sector, region) pairs directly linked to deforestation. By performing this analysis for each pair in the unique list of (sector, region) pairs for each portfolio company, we derive an indirect exposure score per company.

Figure 10 takes the companies in the MSCI ACWI and plots the median IO scores (red lines), where the middle 50% of the scores lie (box), as well as the 'whiskers' which extend to the furthest data points within the 1.5 times from the edges of the box to 1.5 times the smallest and largest within 1.5 times the interquartile range (third quartile minus the first quartile)[13]. For the sake of this report, it is sufficient to understand that the figure shows:

A that on average, the IO score corresponds with the intuition that for downstream sectors that rely indirectly on upstream sectors heavily exposed to deforestation (i.e. the packaged food sector which relies on agricultural production) the score is high, while for those that do not (i.e. the semiconductor industry), the score is low

B that there is significant heterogeneity within the scores for each sector (e.g. there is a large range of IO scores assigned to different companies in the packaged food sector). This shows, nicely, that the IO score is company and location specific. The connection to upstream high-risk (sector, region) pairs depends on where the downstream company operates, as well as whether the company might have part of its revenue/subsidiaries/ assets (see Step 0) associated with other sectors than its primary sector classification.

Deforestation Exposure of Financial Institutions via Forest & Finance [GC Decision Tree 1, Step 1]

Forest-risk sectors such as agricultural production, logging and mining require significant capital investments. Despite industry pledges, commercial finance continues to flow to companies performing high-risk activities in high-risk regions. Among the many ways such financing can be structured are direct loans, as well as the purchasing of bonds and equity (Forest & Finance, 2023). As described previously, deriving a single indirect flag for the financial sector is not only inaccurate, but flagging the entire

[12] https://onlinelibrary.wiley.com/doi/10.1111/jiec.12715

[13] For more information on how to interpret boxplots, see Wikipedia, for example (link).

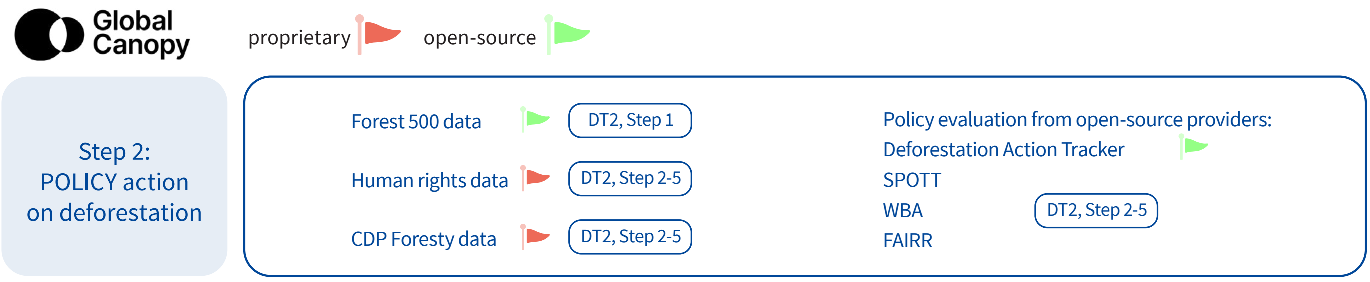
financial sector as 'high-risk' without further screening is also not an actionable way forward. We therefore propose to rely on the Forest and Finance dataset (in addition to Trase, which is covered separately below) to derive company-specific scores/flags as an alternative to the indirect sectoral flags. Forest and Finance assesses the financing by over 300 companies whose operations are likely to impact tropical forests directly given that they operate in high-risk supply chains (i.e. beef, soy, etc) and regions (Brazil, Indonesia, etc). More detail on the Forest and Finance dataset as well as implementation guidance can be found in **Annex Step 2**.

Trase Data [GC Decision Tree 1, Step 2]

Deforestation is driven in significant part (~40%) by the production of commodities such as beef, soy, palm oil, pulp, paper, and minerals (Global Forest Watch, n.d). The Trase Earth database focusses on these commodity supply chains. It leverages disclosed trade (export) data: commodity type, volume, exporter and destination port (consumer side) and uses a material flow analysis model (SEI-PCS) to link these to specific production regions (Trase, n.d.). This connection allows an estimation of the deforestation associated with a certain volume of a particular commodity tied to a specific exporting legal entity and with a specific consumer destination. Before its decommissioning in 2023, Trase Finance (Global Canopy, SEI, Neural Alpha 2023) used publicly and proprietary datasets (e.g. Refinitiv, GLEIF, annual reports, etc) to map financing and ownership of the traders mapped by Trase Earth. This was of particular value for the analysis highlighted in this report because commodity traders are often not public entities. They are frequently owned as subsidiaries or through joint ventures by larger (potentially public) entities or financed by them through equity, bonds, or loans. Analysing Trase Earth directly, without tracing the financing and ownership, risks undercounting the true commodity-driven deforestation exposure of equity portfolios. In our MSCI ACWI pilot, we used the now decommissioned Trase Finance (Global Canopy, SEI, Neural Alpha 2023). Workarounds for future implementation are described in **Annex Step 2**.

Step 2 – Policy Evaluation

Figure 11 - Step 2, Deep Dive



Having analysed the exposure of portfolio companies to deforestation, the user gains an accurate understanding of the likelihood that these companies are involved in deforestation. The next question is what the companies are doing about it. This step operationalises the methodology of Global Canopy et al.'s (2023) Decision Tree 2, which provides a structured framework for assessing a company's policy actions on deforestation and human rights. This framework captures to what extent the risk associated with deforestation exposure (as measured in Step 1) could be mitigated by progressive policy, and therefore be transitory. Global Canopy et al's (2023) guidance outlines five steps to evaluate a portfolio company's policy strength, which we split into three sub-steps: leveraging F500 data (Step I), evaluating the existence of a deforestation and human rights policy (Step II), and evaluating the strength of the policy regime (bundling Steps III-V). Note that this step might overlap with existing processes for controversy screening regarding human right violations and/or environmental controversies. If such processes are in place, we suggest using these synergies.

Leveraging F500 Data [GC Decision Tree 2, Step I]

The Forest 500 list includes 350 high-impact companies and 150 financial institutions. As this database is updated annually, focuses on high-impact companies, and includes a comprehensive assessment of companies' deforestation policies, it is used in this prominent place.

Forest 500 database. The Forest 500 list compiles publicly accessible information on companies and evaluates their ambitions and commitments to mitigate deforestation within their supply chains and tackle related human rights concerns. The assessment provides detailed insights into each company's approach across different commodities, enabling a comprehensive evaluation of a company's overall performance as well as the ambition, scope, and strength of its deforestation and human rights commitments. A string-matching algorithm for matching portfolio companies to the Forest 500 database based on company names can be found in the online code repository[14].

Existence of Human Rights and Deforestation Policies [GC Decision Tree 2, Step II]

This step generally assesses whether a company has a policy on deforestation and human rights. To do this, we collect data from the various sources below and code it into binary indicators. This provides a comprehensive overview of whether human rights and deforestation policies are in place.

CDP Forests Questionnaire. The CDP Forests questionnaire from 2023 contains detailed questions on corporate deforestation policies for a sample of

[14] One limitation is the ability to link the information to other data sources and financial data as no unique identifier such as ISIN, RIC or Ticker are provided (one can get access to an ISIN-linked version through purchasing a license to ForestIQ), see link.

544 companies (327 of which are publicly listed). This proprietary survey is comprehensive, covering topics from general deforestation policies to management incentives and detailed queries about production, consumption, and monitoring practices. The inclusion of ISINs in the dataset makes the data integration straightforward. We encode certain questions relevant to assess a company's policies. Our process for utilising this data is outlined in **Annex Step 2**

Human rights data from ESG data providers. Beyond data on environmental and deforestation policies, it is crucial to assess how companies address human rights issues more broadly, deforestation-related human rights abuses particularly. For this purpose, we primarily use Refinitiv to extract variables on the existence of a human rights policy, as well as an overall score of the effectiveness of human rights policies. Additional ESG data providers like MSCI, Bloomberg, and S&P can offer supplementary or alternative insights.

Human rights and deforestation data from open-source evaluations. In this step of our analysis, we consolidate publicly available information from existing third-party assessments in the field of Nature Finance. These include the World Benchmarking Alliance (which assesses the 380 most influential companies in the food, agriculture, paper and forestry sectors), SPOTT (which assesses 230 companies in the palm oil, timber and rubber sectors), and the FAIRR Protein Producer Index (which assesses the 60 largest protein producers). For financial institutions, we incorporated the Deforestation Action Tracker developed by Global Canopy which assesses over 700 financial institutions on the strength of their policies on deforestation, conversion, and associated human rights abuses. We collect and clean this data from their websites, encrypt it where necessary and assign it to the relevant companies. Although these assessments tend to focus on high-impact companies, resulting in limited data availability, the information gleaned is highly relevant.

Positive flags. To add another perspective, we also looked at forward-looking commitments, i.e. whether companies are signatories to initiatives such as the Science Based Targets Initiative (SBTI) or early adopters of the TNFD. Data on these initiatives is often scarce and may be subject to selection bias, as they tend to include companies that are actively pursuing sustainability goals. Given these limitations, we treat companies' participation in these initiatives as a positive signal to the market, indicating that they have to some extent implemented related policies. These indicators are not incorporated into the firm-level aggregates (see Step 3) but are included in the open source Excel as an additional layer of information.

Strength of the Policy Regime [GC Decision Tree 2, Step III-V]

Having assessed the mere existence of human rights and deforestation policies, we then assess the strength of the related policies. Rather than just looking at binary indicators, we use the existing assessments mentioned above. For example, SPOTT evaluates commodity-related policies and assigns an overall score. Based on individual thresholds, we classify companies into low, medium or high-risk buckets. For a detailed description of this process, please refer to Step 2 in the Appendix.

STEP 3 – AGGREGATION & DISCLOSURE KPIS

After collecting, cleaning, and reconciling the various datasets, tools and models detailed in the previous chapters, the user receives an Excel file with the various indicators, which essentially allows practitioners to do their own screening, weighting, and comparison with different universes. In addition, we also include suggestions for aggregating the results at the company level. See the online code repository for an illustration based on open-source data (link). Recognising that different users have different needs; we discuss two options in this chapter. The first follows the guidance by Global Canopy et al (2023) and classifies companies into low-medium-high buckets for exposure and policy assessment and overlays the two dimensions. The second combines the collected indicators into a numerical score.

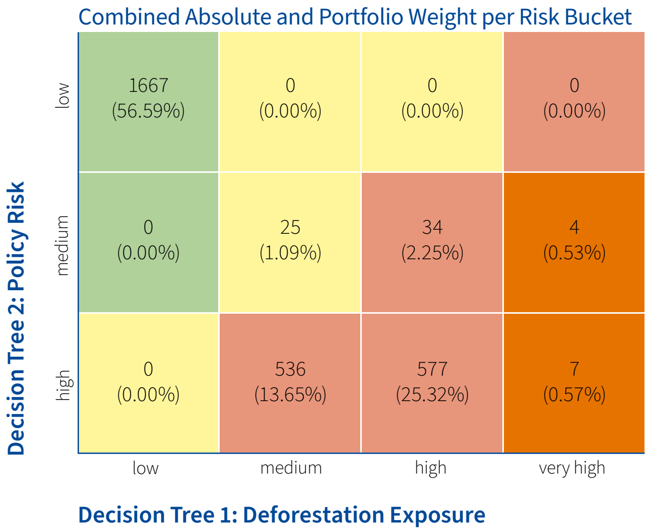
Aggregation via High-Medium-Low Buckets (Categorical)

Inspired by Global Canopy et al (2023), we have implemented a set of decision rules that group companies into specific low, medium, and high buckets for both exposure (Step 1) and policy assessment (Step 2). In this method, data points are evaluated sequentially for each company, and at each step a company is either assigned to a risk bucket (low, medium, high) or moved to the next step in the decision tree. For example, a company is placed in the 'high risk' exposure bucket if its supply chain exposure score is above a certain threshold. Note that we have added a 'very high' bucket as part of Decision Tree 1 for companies in high impact sectors with assets in or near deforestation hotspots. Further details on the decision rules and cut-offs we used for each metric can be found in **Appendix Step 3**.

To reflect the fact that both deforestation exposure and the actions companies are taking to address deforestation exposure are important, the matrix below overlays both buckets, with each cell referring to the number of MSCI ACWI companies and the weighted portfolio share. This overlay allows the financial institution to focus on companies in the bottom right-hand corner of the table that warrant further scrutiny or follow-up with engagement questions.

Companies in the bottom right-hand corner are characterised by a high likelihood to be exposed to deforestation (Decision Tree 1) and insufficient deforestation risk management (Decision Tree 2).

Figure 12 - Resulting buckets for the MSCI ACWI



Note that we followed a fairly conservative set of decision rules, resulting in a high number of "high risk" portfolio companies. This could be altered by the user, see for example AP2's take in the section "Using results in practice". Another key flaw of placing companies into buckets is that it does not fully utilise the detailed information available. For example, a company with a Forest 500 score above 60 is placed in the medium policy risk bucket, failing to differentiate between scores of, say, 61 and 85. Therefore, another potential approach is the weighted scoring approach (see below).

Aggregating via a Weighted-Scoring Approach (Numerical)

Alternatively, one could create a numerical score by weighting the individual indicators. This method allows companies to be ranked against each other, enabling financial institutions to integrate it into metric-based strategies. However, a clear disadvantage is the challenge of determining the relative importance of each variable. While our complementary code repository allows users to set these weights, we have provided a default approach (see **Annex Step 3**). For both exposure evaluation (Step 1) and policy assessment (Step 2), we derive a score for company i , which is the weighted average of the individual scores. Each individual score j (such as the score derived from the IO model) receives a specific weight (for example 30%). Both scores are calculated using the following formula:

$$Score_i = \sum_{j=1}^{M} weight_j \times score_{i,j}$$

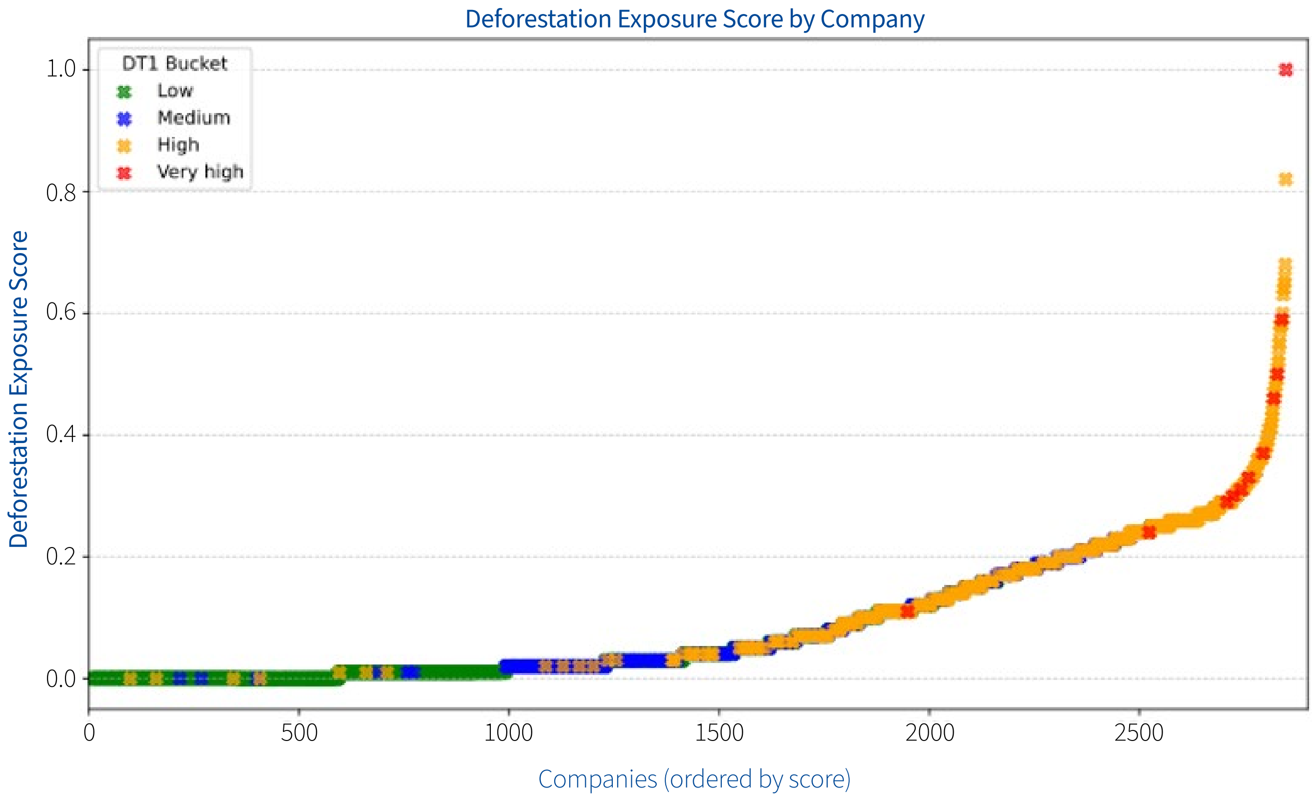
We use different weights for companies and financial institutions. Most variables are also normalised to reflect different units and ranges. The scoring works particularly well for the exposure scores, where we can rely on deep, heterogeneous data (as some companies are scored based on hundreds of location-sector pairs), compared to Decision Tree 2, where all variables refer to the corporate governance level. The DT1 exposure scores therefore produce numerous variations, reflecting the depth of the input data. See **Figure 13** below.

TNFD-Aligned Disclosure

This analysis can be readily used to report on portfolio metrics that align closely with the TNFD. During its launch, the TNFD suggested two core metrics for financial institutions: A) The absolute amount or percentage of the portfolio with exposure to sectors with material dependencies and impacts; B) The absolute amount or percentage of portfolio companies with activities in sensitive locations.

These metrics serve as excellent starting points for exposure-related assessments. We propose extending this by also disclosing policy performance of portfolio companies and the actions taken by the financial institution. Consider the following types of disclosure indicators:

Figure 13 – Deforestation Exposure Scores



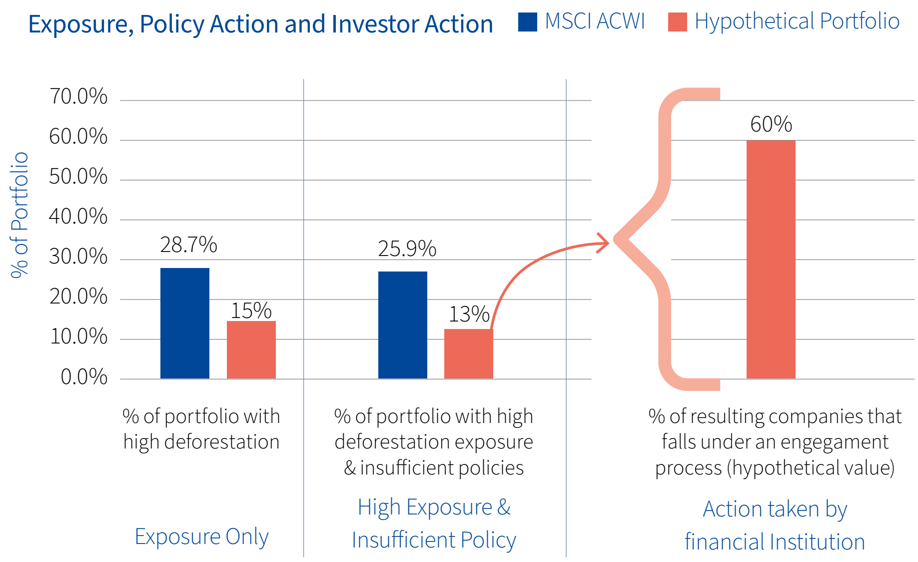
- Exposure-related metrics:
  % of portfolio companies with high deforestation exposure
  % of portfolio companies with high deforestation exposure & sites in proximity to deforestation hotspots
- Exposure & policy performance:
  % of portfolio companies with high/very high deforestation exposure and insufficient policies (overlaying exposure & policy action)
- Action taken by financial institution:
  % of engaged high-risk companies

These metrics can be disclosed and compared to a benchmark. **Figure 14** illustrates these figures for the MSCI ACWI benchmark: 28.7% of the MSCI ACWI (weighted by market capitalisation) fall into the high or very high-risk category following the steps outlined in Step 1. Taking policy measures into account, 25.9% are likely to be exposed and do not have an adequate set of policies on deforestation and human rights. This group of companies could, after validation, become part of an engagement strategy that a financial institution could report on.

This is not intended as a definitive approach but rather to stimulate ongoing discussions about meaningful disclosure metrics.

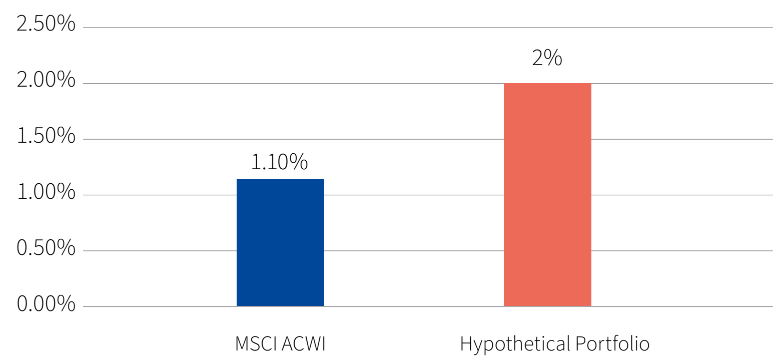
Figure 14 - Portfolio KPIs

Building upon TNFD Core Sector Metric for Financial Institutions #1: Exposure to sectors.



TNFD Core Metric #2: Exposure to sensitive locations

Companies with high impact activities in or near deforestation hotspots



USING THE RESULTS IN PRACTICE (AP2)

AP2 has identified deforestation as a material topic and is, together with 30+ financial institutions globally, a signatory to the Financial Sector Commitment on Eliminating Commodity-driven Deforestation, which was announced in connection with COP26, in 2021. The signatories of the commitment are collaborating in its implementation, in the Finance Sector Deforestation Action (FSDA) initiative. The fund is working towards a portfolio free from commodity-driven deforestation by 2025. and strives to achieve this by identifying companies in its investment portfolio that are linked to high deforestation risk and engaging with these companies as active owners.

To put these words into practice, AP2 has worked with Climate & Company to develop guidelines and open-access data and methodologies to promote deforestation due diligence in practice. This process enables financial institutions to screen and monitor the investment portfolio for deforestation risks. In AP2's view, the dataset developed by Climate & Company has significant advantages over previously available data:

- it can be applied to the entire investment universe.
- sector coverage is broader than previously available data.
- it covers supply chains through the input-output model.
- it includes asset location data linked to sensitive locations (deforestation hotspots).

Applying the Climate & Company methodology and resulting dataset to AP2's portfolio, the Fund combined the different data points for deforestation exposure (see Step 1), to calculate an overall risk rating and assess the deforestation risk of the companies in the portfolio. A percentile-based score was applied to each indicator to be included, all in the range from 1 to -1, ensuring comparable distribution. AP2 further decided to give a higher weight to the supply chain score derived from the IO model, and to combine the indicator on proximity to deforestation hotspots with sector exposure to include assets relevant to deforestation in the analysis.

The analysis resulted in a focus list of 155 companies with high or very high risk, corresponding to about 10% of the Fund's listed equity portfolio. Sectors represented frequently among these companies include food & beverage, metals and mining, pulp & paper, and apparel. The list does not include financial sector companies, which will be addressed separately.

The next step for AP2 was to assess how well the companies on the focus list manage their risk (see Step 2), and as a first indication, a combination of Forest 500 scores and internally performed analysis was used. SBTi commitments or targets set and TNFD Early adopters, both of which are included in the dataset, were viewed as positive indicators.

By combining the two parameters of deforestation exposure and management of risk, AP2's focus list can be mapped as illustrated in **Figure 15**. The Fund will prioritise companies in the bottom right corner for engagement, aiming to engage with 100% of very high-risk companies by 2025. Divesting from high-risk companies with insufficient management of deforestation risks is seen as a last resort but may also prove necessary.

Figure 15 - Deforestation Exposure (DT Bucket 1)

Deforestation risk Managament (DT Bucket 2)	High	Very High
Strong	2	10
Partial	42	21
Weak	59	21

CONCLUSION & CALL TO ACTION

Call to Action for Financial Institutions

The financial sector has a pivotal role to play in the transition towards a deforestation-free economy. This report is a call to action for financial institutions to implement due diligence processes to assess deforestation risk, describing in detail how data can be collected (Step 0) and how exposure can be assessed (Steps 1 and 2). The methodology outlined in this report, along with the online code repository, forms a prioritisation and management tool, helping practitioners to assess deforestation-related aspects throughout their entire portfolio in a structured, data-based, and cost-efficient manner.
We suggest to:

- **Integrate the methodology into your broader nature-related risk strategy.** View this not as a stand-alone exercise but as a hands-on starting point to assess nature-related risks, with significant synergies to the TNFD's LEAP framework and/or the double materiality assessment that is part of ESRS. For example:
  - **Data Foundation:** Step 0 creates the data foundation for location-specific assessments that are crucial for nature-related risk and impacts assessments[15].
  - **TNFD Core Disclosure Metrics:** Following our methodology, it is easy to derive TNFD core disclosure metrics such as the percentage of portfolio companies in sectors with material dependencies and impacts (or deforestation sectors) or the percentage of high-impact companies with locations in or near nature-sensitive areas (or deforestation hotspots).
- **Use these resources for pre- and post-financing decisions to maximise investor impact.** Performing this analysis purely for disclosure will not lead to a change in real-world parameters. Specific steps must be implemented to make full use of the investor contribution potential, such as[16]:
  - **Pre-financing:** Include into investment analysis and due diligence processes pre-investment. Set and communicate investor expectations relating to deforestation (equity) or negotiate financial covenants that allow debt holders to withdraw financing if a time-bound action plan is not met (debt).
  - **Post-financing:** Engage in dialogues with portfolio companies and track their progress towards deforestation free. Set and communicate a public voting policy. Follow an escalation strategy and file shareholder resolutions, exercise your voting rights, and consider divesting from companies if investor expectations are not met (equity). Alternatively, maintain an ongoing dialogue and actively monitor your clients' progress towards time-bound targets (debt).

More comprehensive actions can be found in the guidance developed by Global Canopy, Neural Alpha and the Stockholm Environment Institute (Global Canopy et al, 2023).

- **Join an investor coalition.**
  - Various financial institutions are already taking action through investor collaborations (such as the Finance Sector Deforestation Action, FSDA) to share lessons learned and, most importantly, to conduct collaborative engagement. Robust scientific evidence suggests that – particularly in public equity markets – engagement is a primary channel for driving real-world change, and collaborative engagement is more effective than individual engagement (Mangot and Koch, 2023; Caldecott et al, 2024).

[15] While this report focusses on deforestation, incorporating further geodata on, for example, water-related aspects, comes at low additional costs.

[16] Extracted from Global Canopy et al (2023)

Furthermore, while we encourage the stand-alone use of the methodology and resources, data and finance experts are available to help with follow-up implementations. This support can deepen the analysis of deforestation exposure or add other environmental dimensions to your nature-related strategy[17].

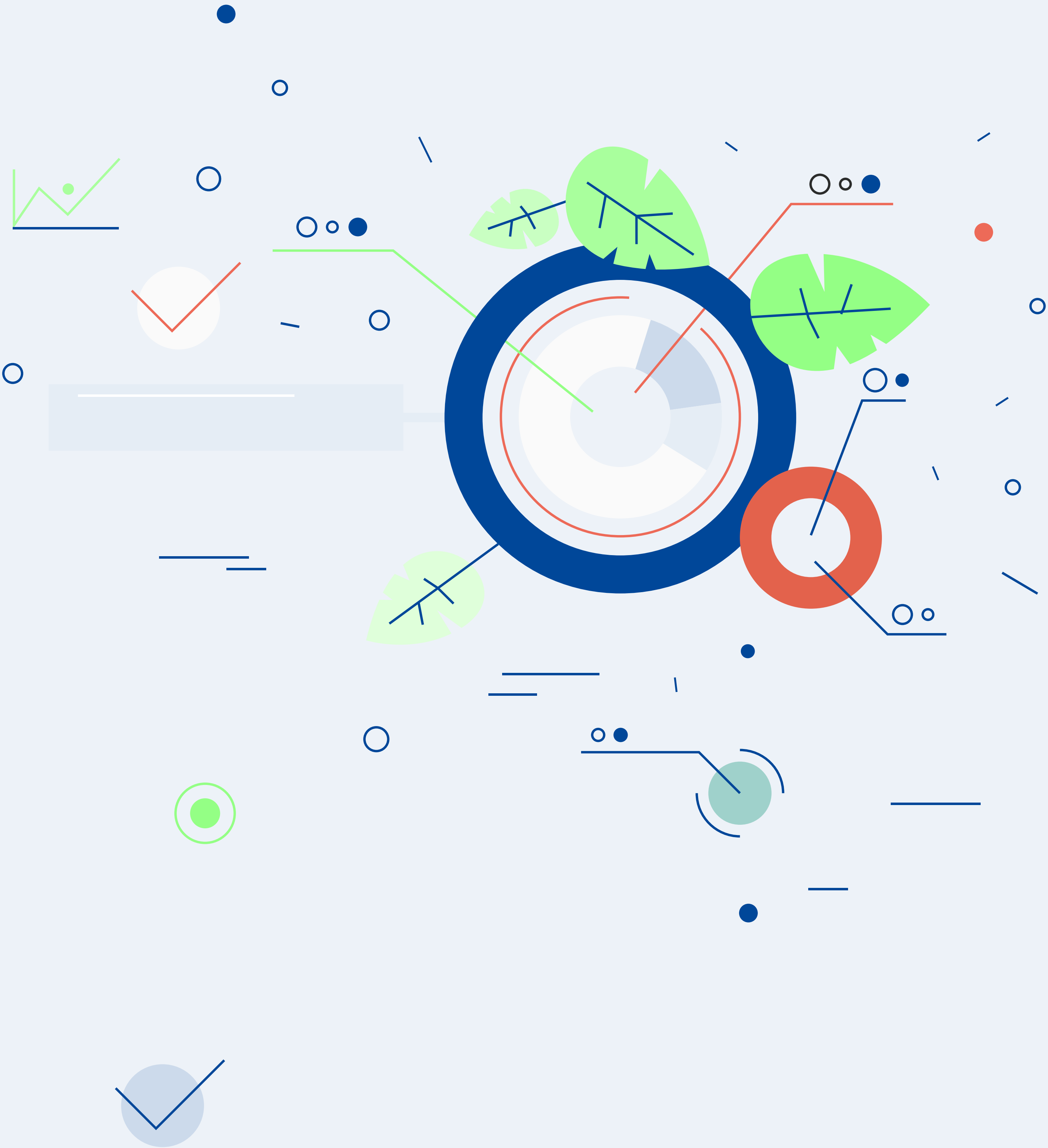
We welcome feedback to improve our methodology, help in adding new datasets, and information on other potential use-cases. For implementation or methodological aspects, you can leave a comment or make a direct pull request on our GitHub repository (a guide on how to do so can be found on the main page). Alternatively, we can be reached by email.

Call to Action for Academics and Civil Society

- **Enhance open-source (location-specific) datasets.** While current open-source datasets serve as a good starting point, the coverage of key aspects can still be improved.
  - **Data on physical assets:** open-source asset-level data linked to companies is still relatively incomplete (see **Figure 5** with physical facilities mapped to approximately 400 companies). High-quality, open-source satellite data, in combination with advances in object classification (machine learning) techniques, will likely increase the automated classification of assets in the future. However, this does not solve the attribution question, i.e. which entity owns/operates the physical asset. The best way of doing so requires additional research and would likely involve a combination of smart engineering and real-life inspection. Relying on NLP techniques to automatically detect information about physical sites could be an interesting avenue. However, until more stringent disclosure requirements are introduced such approaches will not be sufficient.
  - **Policy / controversy data:** The gap between proprietary policy data and open-source alternatives is rapidly closing due to the rise of NLP techniques such as ClimateBERT, as well as more general-purpose foundation AI models such as ChatGPT. We welcome the use of such NLP techniques to analyse nature-related ambitions in corporate annual reports, as well as identify nature-related company-specific controversy from news items in near real-time, and to make such data available open-access.
  - **Supply chain data:** Agricultural commodities are largely driving tropical deforestation. The direct exposure of global equity portfolios is almost non-existent, making it crucial to investigate supply chains. While our methodology used state-of-the-art data, pinpointing exact deforestation exposure for an entire portfolio remains challenging. While input-output models are a good first proxy to cover supply chain aspects of the analysis, there is a great need for more complete open-source data on supply chain relationships. Proprietary data is costly and incomplete. NLP techniques described above could identify suppliers and customers from company websites and reports. Using reverse disclosures one could increase coverage. For example, rather than trying to identify the suppliers to a big corporate, one would look at the customers of smaller suppliers, as these are more likely to mention large corporates to signal successful market access.
  - **Off-the-shelf regulatory data.** To better incorporate (evolving) regulatory risks, better region-specific datasets on evolving nature regulation are needed (as noted also by WWF's terms of references, link). Financial institutions (FIs) making risk assessments of deforestation regulations in specific countries or regions (e.g., EUDR, US Forest Act, Trade Agreements) can combine these assessments with underlying data to determine the most exposed stocks.
- **Leverage recent developments in nature and deforestation data.** We encourage academics to use recent advancements in nature and deforestation data. There is a growing availability of open-source data on companies' nature-related policies as well as granular location data. This data should be used more in research to test the validity of these indicators and metrics, thereby supporting their adoption by practitioners.
- **Investigate financial market implications.** A recent working paper by Bohnet, Fliegel, Tax (2024) found that financial markets increasingly price in companies' deforestation exposure in their long-term asset pricing exercises. By using the methodology described in this report, the authors examined how regulatory events can have severe negative impacts on companies' stock returns. However, there is currently no research looking into other instruments like bonds and loans. Moreover, there is a lack of understanding of how SMEs are exposed to deforestation related policy shocks, and how these affect supply chains.

[17] A non-exhaustive list: Forest IQ, Neural Alpha, Pollination Group. Climate & Company is also about to launch a spin-off to provide data and consultancy services. Do not hesitate to get in touch with the authors to learn more.

climate &
company



General Note on Annex: In this Annex, we aim to clearly outline the technical details, complemented by online resources to facilitate implementation for practitioners. We start with Step 0, which explains how we collected, cleaned and compiled the data necessary for our model to classify companies into deforestation exposure and policy risk groups. Step 1 describes how company-specific deforestation exposure is estimated, while Step 2 focuses on companies' policy evaluation.

We constantly refer to the publicly accessible GitHub repository (link). Please note that the public repository is a restricted version based on open-source data only. Some code snippets have been removed as they rely on proprietary data. The public version is executable and can be modified if users and practitioners wish to include their own datasets. Note that the results when running the public repository will therefore differ from those presented in this report as some data points have been removed. This applies, in particular, to the "input data" (such as the company hierarchy data obtained from ORBIS).

ANNEX STEP 0 (DATA COLLECTION)

Step 0 uses data on company location and industry from open-source asset-level datasets (i.e., asset ownership per location and sector), company structure data (i.e., location and sector of subsidiaries), disaggregated revenue data (i.e., revenue generated per location and sector), and "headquarters information" (i.e., a company's main sector and location). This is collected from three open-source and two proprietary data providers. During the data collection process, a few data cleaning steps and workarounds are required to maintain a large sample size and obtain a usable dataset[18]. We implement different workarounds to obtain a usable dataset. These challenges are already described in detail in WWF's Biodiversity Risk Filter Methodology documentation[19] (see Guidance A, p. 18 ff).

After cleaning every individual dataset, the datasets are merged into one to make use of their complementary elements. As a result, the user obtains a list of location-sector pairs per company that builds the data foundation for our analysis. This subchapter outlines the different components that can be found in the GitHub repository.

Details & Implementation Guidance: Main Portfolio data (incl. headquarter & sector)

First, the user needs to define the main universe to be analysed, which includes basic company level information such as identifiers, but also location and sector information (i.e. the company's headquarter location and primary sector classification). The open source repository uses publicly available data from permid.org for MSCI ACWI companies, while our ("internal") analysis relies on data from Refinitiv. The data generated in this step serves as a starting point for the analysis.

GitHub

Preparing main portfolio data (code)

The open source code repository contains a df_portfolio.xlsx input file which uses data from permid.org, supplemented by a few manual steps. It contains basic identifiers as well as the primary sector classification and country of the companies. This input data can be modified by the user.

As we used data from LSEG/Refinitiv in our analysis, we provide a placeholder script called <prep_refinitiv.py> including a short description explaining which variables could be downloaded. This can also be adapted for data from Bloomberg and other ESG data providers.

Details & Implementation Guidance: Open-Source Asset-Level Data

First, we collect data from Climate Trace, the Spatial Finance Initiative and the Global Energy Monitor. In total, we manage to compile and clean more than 8,000 assets (i.e. physical production sites) and match them to their respective owners. Once we have combined these datasets, including information on the industry and location of each asset, we check whether a unique company identifier is available to link them to portfolio companies. If not, we rely on string matching to link the assets and their owners to the portfolio companies. In addition to the valuable and accurate location information, the additional data on the industry affiliation of the assets allows us to approximate the importance of each location-industry pair per company.

[18] Problems that occur are for example the following: The open-source asset-level data we obtain might include missing values on the production capacity, which is a good proxy to derive the importance and size of the production site. To not lose this information, the missing value could be replaced by the corporate median.

[19] See WWF and Climate & Company (2023). Tackling Biodiversity Risk – A biodiversity risk assessment guide for companies and financial institutions. (link), or WWF Biodiversity Risk Filter Methodology Documentation (link). The first report provides a high-level summary incl. a case study for a subsample of the MSCI ACWI, whereas the second report contains detailed descriptions on the data collection process for financial institutions (see Guidance A).

Other data providers: Other or further data providers can obviously be used for this assessment. GRESB (Asset Impact) for example is a commercial provider that consolidates data from different sources. Other open-source alternatives include the EU-ETS (with comprehensive site-level data on GHG-intensive industries in Europe), the Global Power Plant Database from WRI (with 35,000 power plants from 167 countries), the Global Tailings Portal (with data on mine sites and tailing storage facilities), among others.

Compiling open-source asset-level data (Code)

Our GitHub repository contains the module <generate_combine_asset_data.py> , which compiles and cleans data from the three open-source datasets Climate Trace, Spatial Finance Initiative, and Global Energy Monitor (see generate_asset_level_GEM.py; generate_asset_level_SFI.py; generate_asset_level_climate_trace.py). Aside from incorporating updates, the module does not require any user input and can either be run as it is, if needed certain assumptions about data cleaning and labelling can be changed. The output is an .xlsx file (see next cell).

Explanation of open-source asset-level data (ready-to-use .xlsx file)

The above module produces an .xlsx file. It contains the links to the parent company (either by name or company identifier), the sector allocation (following the NACE sector taxonomy) and an estimate of the importance per asset. It could be used as a starting point for feeding location-specific data into risk and sustainability management processes. However, a precise company identifier (such as ISIN or SEDOL) is not always available and mapping the physical production facilities to portfolio companies therefore relies on text-based string matching. The module described below contains code snippets that could be recycled to do the job.

Mapping asset-level data to portfolio companies (Code)

Although the data collected contains links to the parent company, in most cases only the name is given, and no unique identifier is provided. Therefore, we used a text-based matching. As input, the user must provide a list of company names based on the underlying portfolio and the algorithm returns the mapped universe. Note that we have worked with conservative assumptions to avoid false positives and that the results could arguably be improved with manual work.

Details & Implementation Guidance: Corporate Structure and Subsidiaries

Second, we rely only on ORBIS where we focus on substantial relations with at least 50.01% of ownership. ORBIS includes millions of companies with a high coverage of SMEs. As this report exemplifies the analysis for the MSCI ACWI, the integration of corporate hierarchy and ownership relations is straight forward via ISINs as the unique identifier. Even though for certain variables ORBIS is prone to missing data, basic information on location-industry is available or can be easily imputed. Moreover, the underlying data from ORBIS can be static or dynamic depending on the exact licensing agreement and data collection. Lastly, depending on your license you might be able to retrieve the data in a more practical format, a comprehensive description can be found in the online appendix by Kalemli-Özcan et al. (2023). For this report we only rely on the most recent ownership data from ORBIS collected via the WRDS (Wharton Research Data Services) platform

Other data providers: Other or further data providers can be used as well such as Factset's Data Management Solution, Bloomberg's Corporate Structure data, or data retrieved via Refinitiv..

Compiling corporate structure and subsidiary data (Code)

The module <prep_hierarchy_data_nace.py> compiles and cleans data from a full download of all first level subsidiary links in ORBIS. By relying on unique tuples of all first level owner-subsidiary links, the module connects the company IDs (ISIN, Bureau van Dijk ID) of all stocks in the portfolio to their respective subsidiaries. This process is restricted to ownership links above a user-defined threshold (50.01%). Next, the module iteratively starts connecting the portfolio companies with their direct subsidiaries before then connecting these level-1 subsidiaries with their subsidiaries. Thereby, the iterative process goes down level by level in the hierarchy structure until no ownership link can be found. As a result, the output csv file contains information about all the directly or indirectly linked subsidiaries of a portfolio company and their respective level in the hierarchy structure. As described in the cell below, we use this information to collect additional information on these subsidiary links.

Note that this file is not part of the online open-source repository as it is based on proprietary data and is heavily tailored to ORBIS data.

Collecting and imputing information on subsidiaries (Code)

This module cleans and compiles collected information on the location (city, country) of all subsidiaries and parent companies, as well as their industry affiliation (NACE, NAICS, SIC). However, for many subsidiaries we face the issue of missing values. Hence, we first impute the country of a company by leveraging existing information on the city and/or the first two string values of their Bureau van Dijk ID. Second, we leverage existing and freely available industry mappings (SIC-NAICS, SIC-NACE) to fill missing values. Moreover, as we have hundred thousand of subsidiaries in our data frame we create dictionaries of industry mappings. This enables us to impute many missing NACE codes. Finally, our resulting csv file contains information on the location of a company's and its subsidiaries' operations as well as their industry affiliation. As explained in the cell below, we utilise this information to derive the importance of country-sector pairs for the parent company.

Note that this file is not part of the online open-source repository as it is based on proprietary data and is heavily tailored to ORBIS data.

Derive country-sector importance based on corporate structure (Code)

Within the module <prep_hierarchy_data_nace.py>, we aggregate the information about a parent company's subsidiaries' locations and industry affiliations. We integrate a pre-defined weighting that allows the user to define whether the level of a subsidiary should be considered here. For now, we assumed that every subsidiary is equally important meaning that the country-sector information of all direct subsidiaries (level-1 subsidiaries) are as important as e.g. level-5 subsidiaries of a respective portfolio company. Thus, for each portfolio company we simply aggregate the country-sector pairs of its subsidiaries. This results in a ratio per country-sector pair for each portfolio company serving as a proxy for how important each country-sector pair is. For instance, assume the following structure of company X:

Parent ID	Subsidiary ID	Level	Country	Sector
X	A	1	Brazil	Pulp & Paper
X	B	1	Indonesia	Pulp & Paper
X	C	2	Indonesia	Pulp & Paper

Company X has three subsidiaries with two in level-1 and one in level-2. As we assume they are all equally important unconditionally on their level, we compute the importance ratio for the two country-sector pairs {(Brazil, Pulp & Paper); (Indonesia, Pulp & Paper)}. As two out of three subsidiaries owned by company X are operating in the Pulp & Paper sector in Indonesia this country sector pairs receives an importance proxy of 2/3 with the other country-sector pair (Brazil, Pulp & Paper) receiving 1/3. One limitation in our data source is that we sometimes get multiple different industry classifications for the same subsidiary. For now, we do not treat them differently and aggregate them as they were two separate subsidiaries.

Note that this file is not part of the online open-source repository as it is based on proprietary data and is heavily tailored to ORBIS data.

Derive longitude and latitude information

While location-specific information is available for most subsidiaries in terms of city and country, geographical coordinates are needed to check whether a subsidiary is close to a deforestation hotspot. Therefore, the <prep_hierarchy_data_nace.py> module also derives latitude and longitude information. To do this, we rely on two helper functions stored in <utils.py>. The 'city2lonlat' function imputes coordinates from cities, while 'fill_missing_location_values' uses the Google API.

Details & Implementation Guidance: Disaggregated Revenue

Third, we use sectoral and geographical revenue splits per company from LSEG (Refinitiv), which provides information on the distribution of company revenues, helping to understand profit margins and in particular highlighting where monetary value is added. The data has a major limitation: while we know in which countries and industries revenues are generated, the exact distribution per country and industry is unknown (in this case, a homogeneous distribution has to be assumed (Carbon4 Finance, 2017)). This was used by Svartzman et al. (2021) and Carbon4Finance (2017). We use the percentage of total revenues generated in each country-sector pair as the importance weight

Other data providers: Bloomberg or FactSet Revere also have disaggregated revenue data, among others..

GitHub

Cleaning disaggregated revenue data
Our GitHub repository contains the module <prep_refinitiv.py>, which prepares the data downloaded from LSEG. In the last step of the module, we clean and compile the disaggregated revenues by country and sector and generate an .xlsx file, which is then used as described below (see next cell). Note that this file is not part of the online open-source repository as it is based on proprietary data and is heavily tailored to Refinitiv data.

GitHub

Preparing disaggregated revenue data
Next the two modules <prep_disagg_data_sec_rev.py> and <prep_disagg_data_geo_rev.py> clean the disaggregated revenue for a company's sectors and geographical locations respectively. Moreover, data quality checks are deployed ensuring to only use information with sufficient quality. The output of both modules is an .xlsx file. This xlsx. files are than combined in the module <prep_disagg_data_analysis.py> by relying on the homogeneity assumption. For instance, if company X generates 50% revenues in Brazil and we know that 70% of all the company's revenues are generated in the pulp & paper sector, the assumption imposes that 35% of all revenues are generated in this country-sector pair (Brazil – pulp & paper).
Note that this file is not part of the online open-source repository as it is based on proprietary data and is heavily tailored to Refinitiv data.

Details & Implementation Guidance: Combining All Collected Datasets

As a final step, we combine the four outlined datasets into one. This final dataset then contains an average importance weight per country-sector pair, while allowing the user to specify how to treat missing values and whether the information from each source is similarly important in the analysis. As a default, we assume that each data source is similarly important and provide a weighting based on our confidence in each data source. In addition, we define the two options to treat missing values as follows: option 1 would bias the results towards missing values, whereas option 2 would bias the results towards existing data. As a baseline, we chose our preferred option to bias towards existing data, which we will explain in the cells below. Given these two user choices, the code merges the three datasets (asset-level, hierarchy structure, disaggregated revenue) and averages the respective importance weights for each country-sector pair. If for one country-sector pair none of the three datasets provides information for a company, we use the headquarter and main sector as a last resort to construct the weight (of 100%) for this country-sector pair. As a result, a data frame is generated with at least one country-sector pair per company. If information about the importance of multiple country-sector pairs was collected this information is averaged ensuring that per company all importance weights of country-sector pairs sum up to 1. As described below, in each step and sub-step this information can be leveraged by country, sector or country-sector pair to attribute specific exposure and policy information per company in the most accurate way. The more information is available the more precise these exposure and policy attributions can be.

Combining all collect datasets (Code)

Our GitHub repository contains the module <prep_weighted_sector_region_pairs.py>, which takes the four datasets: asset-level, hierarchy structure, disaggregated revenue, as well as headquarter and sector information and combines them in a structured way.

The structure depends on whether the user wants to bias the results more towards missing data, while always including information on headquarters and sector information. Alternatively, the user can choose a bias towards existing data and rely on headquarter-sector information only if none of the other three data sources provides information. In addition, the user can specify how important each data source is, which is set in the <run_dt1.py> module discussed in Step 3 of the appendix. The idea is that the module <prep_weighted_sector_region_pairs.py> takes this user confidence in each data source and scales the importance weights provided by each dataset by the respective confidence. As a default option, we rely on the existing data bias approach and assume equal confidence in each of the four datasets. In the next cell we describe the intuition behind the two bias options, including an example.

Explanation of the bias options (Code)

A bias towards missing data would mean that for each country-sector pair per company we average the importance weights from each dataset, treating all NAs as zero. This would imply that if company X only has one importance weight for a country-sector pair from the asset-level data, then all missing information with respect to the subsidiary data is valid in the sense that the company has no subsidiaries in that country-sector pair. Thus, any missing information would be assumed to be true and not because the data providers a user relies on are not always 100% accurate. In other words, if you believe that the underlying data you are using does not (almost) perfectly capture the real world, then the bias towards missing data might be a bad choice. In contrast, the bias towards existing data does not consider the fact that a missing importance weight may simply mean that this country-sector pair is not actually important. Here, the module first checks for each company which datasets (asset level, hierarchy structure, disaggregated turnover) provide at least one importance weight. This ensures that missing values from a data source with no information for that particular company are not used and treated as zeros in the averaging process. While it may still be the case that missing information for a particular company's country/sector pair is incorrectly treated as a zero, this reduces the bias towards missing data and thus increases the bias towards existing data. Some may wonder why we call it a bias towards existing data, but we believe that by not accounting for the possibility that a company has no assets, we are overweighting information from the data sources from which we have information.

Resulting output dataframe (Code)

Given similar confidence in each data source and choosing a bias towards existing data, for company X we would get the following country-sector weights if we collected information for at least one country-sector from every data source:

company	country-sector	weight_asset	weight_hierarchy	weight_revenue	weight_hq_sector	weight_final
X	Brazil - 1	0.8	0	0.5	1	1.3/3 = 0.43
X	Brazil - 5	0.2	0.3	0.2	0	0.7/3 = 0.23
X	Germany - 1	0	0.7	0.3	0	1/3 = 0.33

As shown in the table above, there exists information from all four data sources for company X. As the the weight_hq_sector is only used as a last resort the weight_final is relying solely on the asset, hierarchy, and revenue weights. The module < prep_weighted_sector_region_pairs.py> takes the average of the three weights per country-sector pair. Because no subsidiaries were linked for sector 1 in Brazil a weight of zero is assigned and included when averaging the weights for this country-sector pair. This results in comparatively high final weight of 0.43 since most of company X's assets and revenues are associated with this sector. As described in Step 1 for the sector flags, assume that sector 1 is but sector 5 is not a high-risk sector (sector risk = 1), then company X receives a final sector score of 0.76 = 0.43 * 1 + 0.33 * 1 + 0.23 * 0.

ANNEX STEP 1 (DEFORESTATION EXPOSURE)

Step 1 compiles data from various sources that shed light on the deforestation exposure of companies. The following modules can be used in conjunction or independently.

Details & Implementation Guidance: Direct and Indirect Sector Flags

This step computes sectoral flags. While it is clearly an oversimplification, we can broadly divide economic activities as being 'not linked to deforestation', 'directly linked to deforestation' or 'indirectly linked to deforestation' (i.e. those that have a high negative impact potential via their supply chain). To reflect this distinction, we manually integrate information on industry specific exposure to deforestation into a joint excel file and manually distinguish between direct impact and/or indirect impact. This classification is based primarily on a finance sector roadmap on commodity-driven deforestation published by Global Canopy (2021), complemented by information from a CSDDD draft by the European Commission (2022) which identified a set of high-impact sectors. We combine the sources into a single 'flag_direct' and a 'flag_indirect' binary variable in a conservative way, i.e. if either Global Canopy or the CSDDD draft indicates the industry as likely being highly exposed it receives a positive direct/indirect flag respectfully. NB: as described in the main report, we make an important exception for the financial sector, specifically for GICS sector codes 40101010 (Diversified Banks) and 40101015 (Regional Banks) which are flagged by the Global Canopy (2021) report, but which still receive a negative flag (see the excerpt in the table below), as we treat financial institution deforestation exposure separately.

List of sector codes with high-impact flags (ready-to-use .xlsx file)

The Excel file <sector_flags_direct_indirect.xlsx> (GitHub filepath: ./data/input/) contains a list of the GICS sector taxonomy and 1/0 indications on whether the sector has been flagged by Global Canopy. Additionally, based on descriptions in the CSDDD draft we manually flag industries. Lastly, if either the Global Canopy flag (gcp_flag) or the CSDDD flag (csddd_flag) equals 1, we decide whether this sector is directly and/or indirectly exposed to deforestation (as highlighted in blue). An excel file snippet is shown below.

lvl4_code	lvl4_sector	gcp_flag	csddd_flag	flag_direct	flag_indirect
10101010	**Oil & Gas Drilling**	0	1	1	0
10101020	Oil & Gas Equipment & Services	0	0	0	0
10102010	**Integrated Oil & Gas**	0	1	1	0
10102020	**Oil & Gas Exploration & Production**	0	1	1	0
10102030	Oil & Gas Refining & Marketing	0	0	0	0
10102040	Oil & Gas Storage & Transportation	0	0	0	0
10102050	**Coal & Consumable Fuels**	1	1	1	0
15101010	Commodity Chemicals	0	NA	0	0
...	...	...	...	...	...
...	...	...	...	...	...
40101010	Diversified Banks	1	NA	0	0
40101015	Regional Banks	1	NA	0	0
...	...	...	...	...	...

Details & Implementation Guidance: Direct Deforestation Exposure Data

In this step we use the estimates of commodity-driven deforestation from the DeDuCE model (Singh & Persson 2024), which integrates remote sensing data on forest loss with agricultural statistics. DeDuCE attributes deforestation to cropland, pasture, and forest plantation expansions and links this to the commodities produced on these lands. It also includes data on the deforestation impact of agricultural and forestry commodity production, trade, and consumption by country and year from 2005-2022. DeDuCE improves upon the Pendrill et al. dataset by incorporating more precise spatial data and extending the geographical and temporal scope, providing estimates for 176 countries and 184 commodities. The commodities that are considered are broadly classified according to the FAO format (except for providing additional detail concerning types of forest plantation). We take the 2018 data to align it with the version of the Exiobase input-output model we use (see the Annex section on the IO model). In the future it would make sense to update the analysis to rely on more recent data. There are several steps that we take to make the data most useful for our context. First, we remove the region-sector pairs with zero of extremely small amounts of deforestation associated (the user can set this threshold; we set it to 0.001 Ha). This yields several thousands of pairs (7634 when using amortised data, 4503 when using unamortised data). Next, we consider the fact that the deforestation attribution data approximates the number of hectares [Ha] associated with the entire region (ISO)-sector (NACE) pair in absolute terms. However, what we want to measure is how much deforestation is associated with a certain amount of revenue earned in each of the region-sector pairs (i.e. a measure with units of [Ha/dollar]. To do so, we need to estimate the economic significance (in dollars) of each of the attribution data ISO-NACE pairs. The most obvious starting point is the Value of Agricultural Production FAOSTAT database which combines primary commodities' production data with producer prices. In most cases this database directly yields a size estimate for the region-sector pair. In the cases where this is not the case, we estimate the size of the region-sector pair using one of the following (in order of preference):

1. In the case of 'Cattle Meat: 1806', we take the sum over all the relevant FAO item codes.
2. In the case of 'Leather: 919' and 'Forest Plantation: 6716', as these are not in the FAOSTAT database, we take these to be a fixed percentage of the total GDP (2018) for the specific region-sector pair (we chose 0.25% and 0.5% respectively, but they can be adapted by the user).
3. In the case where FAOSTAT information on a specific region-sector pair is missing, but it does contain information on similar commodities, we take the average for that region over all the other FAOSTAT sectors that belong to a similar commodity group (see the 'lookup_adapted' sheet of the 'deduce_lookup.xlsx' file).
4. In the case where there is information on commodity production (FAOSTAT) as well as producer prices (FAOSTAT) we multiply the values to get the total size in absolute dollar terms.
5. In the case where production data is available, but no information on prices for that specific region, we take the average price for the commodity over all the other regions that do have producer prices available. In cases where this still does not yield a price estimate, we perform manual research.
6. In the cases where there is price information available for the region-sector pair, but no production amount. We perform manual research.
7. Finally, in the cases where none of the above worked, we either found FAOSTAT values for 2017, or perform manual research.

Now that we have both the deforestation attribution (in Ha) as well as the economic significance (in millions of dollars), we map the deforestation attribution data to the region-sector pairs generated by Step 0. To do so, the attribution data, as well as the sizes are separately aggregated from the ISO-FAO level to the ISO-NACE level using a manually constructed many-to-one mapping between FAO and NACE (Rev2) codes. This mapping can be found in the repository (nace_to_fao.xlsx). Finally, we divide the attribution data by the size of the region-sector pair to get the relative measure of deforestation per million dollars.

GitHub

Overlay with direct deforestation attribution data
Our GitHub repository contains the module <generate_direct_attribution_filter.py>, which gets called by the <apply_direct_attribution.py> in case the direct deforestation footprints for all the relevant region-sector pairs are not yet computed.

Details & Implementation Guidance: Proximity to Deforestation Hotspots

Aside from a list of region-sector pairs, Step 0 yields precise coordinates for open-source asset level data as well as subsidiaries (corporate structure). With this information we can check whether corporate assets or subsidiaries are in proximity to deforestation hotspots. Note that in the context of EU regulation this step has close synergies with the Principal Adverse Impact indicator #7 (“the share of investments in companies with sites or operations located in or near to biodiversity-sensitive areas, which negatively affect those areas”) as defined under the EU’s SFDR regulation.

To identify deforestation hotspots, we rely on data from Global Forest Watch. The emerging hot spots data set identifies the most significant clusters of primary (tropical) forest loss between 2002 and 2023 (and likely updated indefinitely into the future). A “hot spot” is an area with statistically significant clustering of forest loss, indicating underlying spatial processes rather than random events. The categories of hot spots include New, Sporadic, Intensifying, Persistent, and Diminishing, each describing the temporal pattern and intensity of forest loss clustering. The analysis uses annual tree cover loss data from Hansen et al. (2013), primary forest extent data from Turubanova et al. (2018), and the ESRI ArcGIS Emerging Hot Spot Analysis tool (Global Forest Watch, 2023).

In our script we load the 2023 GFW data on deforestation hotspots between 2002-2023 and asset location data: either disaggregated assets or subsidiaries. The script calculates distances between each asset and the nearest hotspot polygon using geospatial tools. The assets that lie within a distance threshold from the hotspot are counted (i.e. “5 assets of company A are in proximity to deforestation hotspots”), and the count is turned into a simple metric by multiplying with an impact weighting (IMPACT_DICT) based on the type of hotspot: Diminishing (1.0), Intensifying (3.0), New (2.0), Persistent (3.0), and Sporadic (1.0). These weights likely reflect the severity and consistency of deforestation impacts, with intensifying and persistent hotspots considered more critical.

GitHub

Overlay with deforestation hotspot data

Our GitHub repository contains the module <prep_hotspot.py>, which uses location specific data from subsidiaries and assets owned by companies and returns the number of assets in proximity to hotspots as well as impact scores.

The user can define the threshold of “proximity” which we define as 50 km by default. If an asset is within the threshold distance of a hotspot, it is added to the count and receives the corresponding impact weight; otherwise, it receives a score of zero. This process is repeated for each asset.

Details & Implementation Guidance: Environmental Controversies

Depending on the users licensing agreements any additional information from proprietary data providers (RepRisk, MSCI, Bloomberg, LSEG, ...) can be utilised to get a more comprehensive understanding about general as well as ESG or even deforestation related controversies. Through LSEG, we collect information on environmental controversies including the number of (most) recent environmental controversies.

GitHub

Preparing controversy data from LSEG / Refinitiv

This module prepares two environmentally related controversy variables. One of them states the number of controversies in the past while the other states the number of most recent controversies (in the last fiscal year). As the next and last step in this module, ISINs are used to easily merge this information to portfolio companies.

Note that this file is not part of the online open-source repository as it is based on proprietary data and is heavily tailored to Refinitiv data. Instead, the repository contains a placeholder called prep_esg_controversies.py.

Details & Implementation Guidance: Indirect Deforestation Exposure via Input-Output Modelling

To approximate indirect deforestation exposure in the absence of complete and reliable supply chain data, we rely on the same DeDuCe deforestation attribution data as above, and 'propagate' it through the economy using the environmentally extended multi-regional input-output model (MRIO) EXIOBASE3. Input-output (IO) models capture the relationships between different region-sector pairs of an economy, detailing how the output from one region-sector is used as input in another region-sector to produce its output. This framework quantifies the interdependencies within an economy, showing the flow of goods and services between industries and their contribution to the overall economic activity. IO models are constructed using national economic data and trade statics, typically compiled from a.o. government statistical agencies, which provide detailed information on production, consumption, and trade across sectors. EXIOBASE3 is an advanced, environmentally extended multi-regional input-output (MRIO) model that balances detailed regional and sectoral information. It provides extensive detail on sectors that drive environmental pressures, such as agricultural production, making it especially suitable for comprehensively analysing global supply chains and their environmental impacts. For more technical details on exactly how MRIOs are constructed, please refer to the original EXIOBASE3 publications and references therein.

We start by taking the DeDuCe attribution data and aggregating the attribution data from the ISO-FAO level to the EXIOBASE_region-EXIOBASE_sector level using both a many-to-one mapping between ISO and EXIOBASE_region, as well as a many-to-one mapping between FAO and EXIOBASE_sector (please get in touch with the main authors if this mapping is of interest to you; i.e. if you want to calculate the aggregated attribution data yourself). The resulting aggregated attribution data can be found as "UpdatedAttributionData.csv". We then calculate the input-output relationships using the Leontief inverse and compute the deforestation footprint per unit output for each region-sector pair. The Leontief inverse is calculated from the direct requirements matrix (A), which represents the inter-industry coefficients showing the input required from each region-sector to produce one unit of output in another region-sector. The Leontief inverse accounts for both direct and indirect economic interactions. The script then uses the total output vector (x) to normalise the deforestation data and compute the deforestation intensities (s), representing hectares of deforestation per unit of output. Finally, the deforestation intensities are multiplied by the Leontief inverse to obtain the indirect deforestation footprint, which reflects the cumulative deforestation impact throughout the entire supply chain and therefore account for indirect effects.

GitHub

Estimating indirect deforestation using EXIOBASE3
Our GitHub repository contains the module <generate_supply_chain_filter.py>, which gets called by the <apply_supply_chain_filter.py> in case the indirect deforestation footprints for all the relevant region-sector pairs are not yet computed.

Details & Implementation Guidance: Deforestation Exposure of Financial Institutions via Forest & Finance

The Forest & Finance dataset meticulously tracks financial flows from financial institutions to around 300 upstream and midstream companies operating in high-risk sectors and regions. It focuses on companies involved in the supply chains of beef, palm oil, pulp & paper, rubber, soy, and tropical timber, primarily in Southeast Asia, Brazil, and Central and West Africa. The companies are selected based on their size, operational land area, and the availability of financing information. This dataset is the result of extensive research by a coalition of organisations, including the Rainforest Action Network, TuK INDONESIA, Profundo, Amazon Watch, Repórter Brasil, and BankTrack.

The methodology applies segment and geographical adjusters and differentiates between various types of funding, including historical data on bond issuance, share issuance, direct corporate loans, and revolving credit facilities. It also includes point-in-time assessments for shareholding and bondholding (Warmerdam, 2020). By linking companies to their received funding volumes, the dataset provides a useful tool for approximating the exposure of financial institutions.

Processing and integration Forest & Finance data

The generate_forest_finance_scores.py file processes the raw Forest & Finance dataset. The script calculates total financing amounts by banks over a specified number of years and handles outliers. The user can specify the year of interest and truncation period.

The apply_forest_and_finance.py file attaches the financing by financial institution to specified universe of portfolio companies via string matching, and creates high-impact flags if the financing amount is above a specified threshold.

Details & Implementation Guidance: Trase (Earth)

As mentioned above, our MSCI ACWI pilot has utilised the now decommissioned Trase Finance (Global Canopy, SEI, Neural Alpha 2023). Currently, we see two ways to address this issue and make use of the publicly available Trase Earth data.

First, one could use the datasets provided by Trase Earth directly. Trase has done the heavy lifting and offers free-to-download supply chain data for a range of commodities in 10 countries (Argentina, Bolivia, Brazil, Colombia, Côte d'Ivoire, Ecuador, Ghana, Indonesia, Paraguay, Peru). For each country-commodity pair, the data includes the country of production, the exporting company, the exporting group, the volume, the country of destination, and whether the trade is covered by a corporate zero deforestation commitment, among other values. By compiling all these data sources, users can check whether their portfolio has exposure to companies captured by Trase Earth data and the respective volumes.

Secondly, one could re-engineer the connection between financing and ownership data and commodity traders in Trase Earth or rely on support from "organisations such as Forest IQ, Responsible Capital, or Climate & Company.

ANNEX STEP 2(POLICY EVALUATION)

Details & Implementation Guidance: Forest 500

One of the most prominent company-specific deforestation risk proxies is the Forest 500 list developed by Global Canopy. This list is made available for free by Global Canopy including detailed and commodity specific assessments of the 350 companies and 150 financial institutions that are most exposed to deforestation. We manage to match around 100 entities of their list based on company and FI names with the portfolio companies. Then we utilise the fact that if a company has been selected in accordance with the Forest 500 selection methodology, it has been identified to have high exposure to forest-risk commodities. Additionally, if available for a company, we use the total score to approximate the strength of a company's deforestation policy and implementation.

GitHub

Compiling and matching the Forest 500 list to portfolio (Code)
Our GitHub repository contains the module <prep_forest500.py>, which combines and prepares the Forest 500 list. First the two lists including the 350 most exposed companies and 150 most exposed financial institutions are combined and merged to user's portfolio companies. The user can specify whether to use a pre-defined mapping or do a fuzzy string matching. The pre-defined mapping is based on fuzzy string matching of the Forest 500 list with the MSCI ACWI combined with manual checks. Hence, if the user is interested in a subset of the MSCI ACWI the pre-defined mapping should be used. Otherwise, we suggest to use the fuzzy string matching, or if possible, carry out a stand-alone string matching including manual checks and save it in a similar way than our mapping file <forest500_matches.csv>. Lastly, the module generates a flag if a portfolio company is on the Forest 500 list.

Details & Implementation Guidance: Open-Source Data on Human Rights/ Deforestation Policy

We leverage several open-source datasets from reporting initiatives such as the World Benchmarking Alliance, SPOTT, SBTN, the FAIRR Protein Producer Index, the Food Emissions 50 dataset, or the Deforestation Action Tracker. Moreover, we try to rely on other open-source information which can be interpreted proactive step taken by companies regarding their nature/deforestation policies (e.g. TNFD early adopter). The aim is to first assess whether a company has a deforestation and/or human rights policy in place, before we approximate the strength of the policy. Each dataset is either downloaded from the respective website or directly web scrapped. Then the data frames are cleaned, and variables explaining the existence and/or strength of a respective policy are identified. Finally, every data frame is merged with the portfolio companies based on any unique company ID provided by the initiative or if no ID is available based on string matching..

GitHub

Generating data frames from open-source data
Our GitHub repository contains multiple modules <generate_FAIRR.py>, <generate_SBTN.py>, <generate_SPOTT.py>, <generate_food_emissions_50.py>, <generate_deforestation_action_tracker.py>, and <generate_WBA.py> which clean, subset and merge the data from these initiatives to the user's portfolio companies. Each module is structured similarly. First, the most informative variables are manually identified to assess the existence and/or strength of human rights and deforestation policies. After the data is subset, company identifiers like ISIN, SEDOL, ticker or LEI are used to merge it with portfolio companies. In contrast, the module <generate_TNFD.py> web scrapes information which companies are early TNFD adopters. Because no unique company identifier is available the module relies on string matching to link to the respective portfolio companies.
Note that the generate_FAIRR.py file is only a placeholder. Users need to request access to the FAIRR dataset

Details & Implementation Guidance: CDP Forests Questionnaire

To integrate the Forests questionnaire from CDP, we had to identify specific questions that are suited for DT2. After we identified them, we encoded them such that given answers are characterised into a numerical number indicating whether the company's answer can be perceived positively or negatively. This does not rely on any normative framing or a textual analysis as many answers are chosen from a predefined answer set. As seen from the example below, a qualitative choice was made by encoding the answer in a certain way, which the user can change depending on their needs. As of 2022 more than 550 companies submitted their survey answers though CDP. However, there are only around 370 companies with a unique identifier that enables us to easily merge the information to the portfolio companies. Hence, similar to the Forest 500 data the coverage of the CDP survey limits us from scaling this up to the whole portfolio.

Processing and encoding of CDP survey

Our GitHub repository contains the module <generate_cdp.py>, which encodes a specific subset of the CDP questionnaire and the respective answers given by companies. Each answer is linked to a company ID (ISIN, Ticker) which allows the user to link it to the portfolio companies. The underlying CDP data can include multiple ID's, for instance two different ISINs. For now, we only rely on the first ISIN provided. Moreover, questions are often answered by choosing from a pre-defined set of answer options. Hence, we encode them as follows:

The question "Does your organization have a policy that includes forests-related issues?" is transformed in a score between 0 and 1 conditionally on whether there seems to be a policy in place:

1. "Yes, we have a documented forests policy that is publicly available" = 1
2. "Yes, we have a documented forests policy, but it is not publicly available" = 0.8
3. "No, but we plan to develop one within the next two years" = 0.3
4. "No" = 0
5. "Question not applicable" = 0

Note that this file is a placeholder in our online repository. Users need to have access to CDP data.

Details & Implementation Guidance: ESG Data Providers

Depending on the users licensing agreements, any additional information from proprietary data providers (RepRisk, MSCI, Bloomberg, LSEG, ...) can be utilised to get a more comprehensive understanding about deforestation and human rights policies. As we only have access to LSEG, we rely on human rights policy variables that either indicate the existence or strength of a human rights policy. We encode the variable regarding the existence of a policy into a binary variable. Lastly, we rely on the ISIN provided by LSEG as a unique company identifier to merge the variables to the portfolio companies.

Preparing policy data from LSEG

In our analysis, we used Refinitiv data and coded variables that indicate the existence and strength of corporate human rights policies.

Please note that this file has been removed. A placeholder file called <prep_proprietary_policy_vars.py> has been added instead..

ANNEX STEP 3 (AGGREGATION)

Details & Implementation Guidance: Aggregation via High-Medium-Low Buckets (Categorical)

Decision Tree 1 – Deforestation Exposure

To implement Decision Tree 1 as outlined by Global Canopy, Neural Alpha and the Stockholm Environment Institute (2023), we have defined a set of decision rules that classify companies into low, medium, high and very high risk buckets, using the data indicators collected in Step 1 as described above. Note that we have added a ‘very high’ bucket as part of Decision Tree 1 for companies in high impact sectors with assets in or near deforestation hotspots. The script below executes several steps sequentially, making decisions based on binary or numeric variables.

The decision rules can be changed by the user (see user_input.py). Here are some examples:

- **Direct and indirect sector flags:** Companies are classified as high (medium) risk if 75% (50%) of their business activities are in high-risk sectors (directly or indirectly).
- **Forest 500:** Companies are assigned to the high-risk bucket if they are part of the Forest 500 list (as Global Canopy manually identifies the 350 companies with the greatest impact on tropical deforestation).
- **Trase:** Companies are assigned to the High Risk bucket if they have been flagged by the (now discontinued) Trase Finance dataset.
- **Environmental controversies:** Companies are assigned to the high or medium risk bucket if they have been involved in recent or historical environmental controversies. The thresholds can be set by the user in user_input.py.
- **Indirect deforestation exposure through IO modelling:** The IO model provides an exposure score for each company, indicating how likely it is that the company is exposed to deforestation in its supply chain. Companies are considered high risk if the score is above 0.1 and medium risk if the score is above 0.04. These thresholds were derived by looking at the distribution of the MSCI ACWI portfolio.
- **Estimated direct deforestation attribution:** This function uses the most recent deforestation data at country and sector level and assigns it to the respective company pairs. If there is an overlay, there is a score above 0, which is used to assign companies to high risk buckets. Note that for a global equity portfolio such as the MSCI ACWI, “hits” are almost non-existent.
- **Proximity to deforestation hotspots:** If a company has activities in or near deforestation hotspots (the user can define the threshold, which is set at 50 kilometres by default), the company is assigned to the very high risk bucket.

GitHub

Deforestation Exposure: Assigning companies into low, medium, high, very high buckets (Code)
The file apply_decision_tree1_logic.py contains the function that assigns companies into buckets, following a set of thresholds that can be set by the user via user_input.py.

Decision Tree 2 – Policy Action on Deforestation

Decision Tree 2 (DT2) focuses on policies and corporate governance related to deforestation. Global Canopy, Neural Alpha and the Stockholm Environment Institute (2023) outlined five steps: using the policy scores from the Forest 500 dataset (Step 1), analysing the mere existence of a deforestation or human rights policy (Step 2), and analysing the strength and content of deforestation and human rights policies, including their implementation (Steps 3-5). In order to carry out the analysis at scale, we have divided the assessment into the following steps:

- **DT2 – Step 0: Dealing with low-exposure companies**

  - All companies that fall into the "low exposure" bucket of decision tree 1 were directly assigned into the low risk policy bucket, following Global Canopy et al. (2023). This follows the assumption that deforestation policy risk is low if exposure is low.
- **DT2 – Step 1: Leverage Forest 500 data [GC DT2, Step I]**

  - Companies are placed in the medium bucket if the Forest 500 score is above 60 (i.e., company has strong deforestation and human rights policies), and in the high risk bucket if it is below 60 (i.e. weaker policy). Note that this threshold is quite ambitious and could also be changed by the user. For the 2023 assessment year, only 8 out of 350 companies had a score above 60.
- **DT 2 – Step 2: Existence of human rights and deforestation policies [GC DT2, Step II].**

  - We extracted and processed variables from the CDP Forest questionnaire and human rights data from Refinitiv. We also extracted relevant variables from the World Benchmarking Alliance, SPOTT, the FAIRR Protein Producer Index and the Deforestation Action Tracker that indicate the presence - or absence - of deforestation and human rights policies. If a company does not have a deforestation policy or a human rights policy, it falls into the high policy risk category.
  - To give an example, the SPOTT detailed palm oil assessment data includes an indicator on "Sustainable palm oil policy or commitment for all its operations. If this indicator is positive, we consider the company to have a deforestation policy. Further details and variables can be found in the code.
- **DT2 – Step 3: Strength of Policy Regime [GC Decision Tree 2, Step III-IV].**

  - Whereas in Step 2 we only looked at the existence of a policy, here we extracted score variables from the different providers and set thresholds that rewarded good scores or penalised worse scores. Full details can be found in the code.
  - To give an example: the Deforestation Action Tracker (DAT) dataset provides a total score variable that summarises the overarching deforestation policy action of 714 financial institutions. Based on this score, we assess the strength of the deforestation policy regime. If the score is >25 (out of 100), it is "strong". If between 10 and 25, "moderate". If below 10, "weak". For example, 69 out of 714 financial institutions in the DAT sample have a score above 25. Similar rankings are made for other variables. If a company has at least a " moderate " human rights policy and a " moderate " deforestation policy, it is placed in the medium policy risk category.

GitHub

Deforestation Policy Action: Assigning companies into low, medium, high high buckets (Code)
The file apply_decision_tree2_logic.py contains the function that assigns companies into buckets, following a set of categorical and numerical thresholds. The logic and strength of the assessment can be adjusted by the user.

Details & Implementation Guidance: Aggregation via Weighted Scoring (Numerical)

Global Canopy, Neural Alpha, and the Stockholm Environment Institute (2023) outlined the approach for categorising companies into low, medium, and high exposure and policy risk buckets. This method is particularly useful when overlaying the two categories. However, a key flaw is the difficulty in distinguishing different levels of deforestation exposure and policy action. Additionally, categorising companies into "buckets" based on thresholds does not fully utilise the detailed information available. For example, a company with a Forest 500 score above 60 is placed in the medium policy risk bucket, failing to differentiate between scores of 61 and 85. An alternative approach involves weighting all relevant variables. The deforestation exposure score example is illustrated in the figure in Step 3.

For both exposure evaluation (Step 1) and policy assessment (Step 2), we derive a score for company i , which is the weighted average of the individual scores. Each individual score j receives a specific weight. Both scores are calculated using the following formula:

$$Score_i = \sum_{j=1}^{M} weight_j \times score_{i,j}$$

We have deprioritised fine-tuning the policy risk score due to the homogeneity of the data, which leads to similar results with noticeable "jumps". However, as shown in Step 3, the deforestation exposure score (DT1) displays numerous variations, reflecting the depth of the input data.

For the DT1 exposure score, we first normalised most variables to account for the different units and ranges. We also distinguished between corporates and financial institutions. To give a concrete example, the results shown in the **Figure 13** in Step 3 were calculated with the following weights for corporates:

- **Direct and indirect sector flags:**
  - flag_direct_score, i.e., the percentage of the company's business activities attributed to sectors with a high direct exposure to deforestation: 12.5%
  - flag_indirect_score, i.e. the percentage of the company's business activities attributed to sectors with high direct exposure to deforestation: 12.5%
- **Forest 500:** flag_forest500, i.e., whether the company is part of the Forest 500 list (1/0): 5%
- **Environmental controversies:** controversies_recent, i.e., the number of recent environmental controversies: 10%
- **Indirect deforestation exposure through IO modelling:** coIO_model_score, i.e., the estimated supply chain exposure: 30%
- **Trase Finance:** Trase_df_exposure, i.e., the deforestation exposure calculated by the now decommissioned Trase Finance data: 15%
- **Estimated direct deforestation attribution:** Direct_attribution_score, i.e. the score reflecting the overlay with the latest deforestation data based on companies' activities in sectors and regions: 5%
- **Proximity to deforestation hotspots:**
  - asset_impact_assignment_count_subsidiary, i.e. whether companies have subsidiaries are in or near deforestation hotspots: 5%
  - asset_impact_assignment_count_assets, i.e. whether corporate assets derived from asset-level daatsets are located in or near deforestation hotspots: 5%

GitHub

Assigning numerical exposure scores (Code)
The file apply_decision_tree1_logic.py contains the function apply_dt1_weighted_average_approach which normalises a set of variables and calculates the scores. The function takes a weight dictionary as an input (defined by the user).

The file apply_decision_tree2_logic.py contains the function apply_dt2_weighted_average_approach. The function works but results are less heterogenous due to substantially less input data.

REFERENCE LIST

Banque de France (2020): A “Silent Spring” for the Financial System? Exploring Biodiversity-Related Financial Risks in France. Retrieved from https://publications.banque-france.fr/en/silent-spring-financial-system-exploring-biodiversity-related-financial-risks-france

Bohnet, M., Fliegel, P., Tax, Y. (2024). Zooming in on Biodiversity Risk: The Pricing of Deforestation on International Financial Markets. (forthcoming)

Caldecott, B., Clark, A., Harnett, E., & Liu, F. (2024). How sustainable finance creates impact: transmission mechanisms to the real economy. Review of World Economics, 1-33.

Carbon4 Finance (2017): Climate Risk Impact Screening: the methodological guidebook. Retrieved from https://strapipreprod.s3.eu-west-3.amazonaws.com/CRIS_Guidebook_Publicversion_November_2017.pdf

Christiaen (2023): State and Trends of Spatial Finance. Retrieved from https://www.cgfi.ac.uk/wp-content/uploads/2023/03/State-and-Trends-of-Spatial-Finance-2023.pdf

European Commission (2022): COM/2022/71 final, Proposal for a DIRECTIVE OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL on Corporate Sustainability Due Diligence and amending Directive (EU) 2019/1937. Retrieved from https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A52022PC0071

Forest & Finance (2023): Banking on Biodiversity Collapse. Retrieved from https://forestsandfinance.org/wp-content/uploads/2023/12/BOBC_2023_vF.pdf

Global Canopy (2021): Eliminating Commodity-Driven Deforestation - Finance Sector Roadmap. Retrieved from https://guidance.globalcanopy.org/roadmap

Global Canopy (2023): Due diligence towards Deforestation-Free Finance – Guidance for Financial institutions. In partnership with Neural Alpha and Stockholm Environment Institute. Retrieved from https://guidance.globalcanopy.org/further-guidance/due-diligence-towards-deforestation-free-finance/

Global Forest Watch (n.d.): Commodities. Retrieved from https://www.globalforestwatch.org/topics/commodities/

Global Forest Watch (2023): Emerging Hot Spots 2023 [Data set]. Retrieved from https://data.globalforestwatch.org/datasets/gfw::emerging-hot-spots-2023/about

Hansen et al. (2013): High-resolution global maps of 21st-century forest cover change, Science, 342(6160), 850-853. Retrieved from https://doi.org/10.1126/science.1244693

Kalemli-Özcan et al. (2024): How to Construct Nationally Representative Firm-Level Data from the Orbis Global Database: New Facts on SMEs and Aggregate Implications for Industry Concentration, American Economic Journal: Macroeconomics, 16(2), 353-74. Retrieved from https://www.aeaweb.org/articles?id=10.1257/mac.20210113

Mangot, M. & Koch, N. (2023). The Impact Potential Assessment Framework (IPAF) for financial products. 2 Degree Investing Initiative. https://2degrees-investing.org/wp-content/uploads/2023/03/The-Impact-Potential-Assessment-Framework-IPAF-1.pdf

Pendrill et al. (2019): Agricultural and forestry trade drives large share of tropical deforestation emissions. Retrieved from https://www.sciencedirect.com/science/article/pii/S0959378018314365

Singh and Persson (2024): Commodity-driven deforestation, associated carbon emissions and trade 2001-2022. Retrieved from https://zenodo.org/records/10633818

Sood et al. (2022): An assessment of the 25 largest European banks' biodiversity and climate strategies. Retrieved from https://cdn2.assets-servd.host/shareaction-api/production/resources/reports/ShareAction_Banking_Survey_2022-final.pdf

Stadler et al. (2018): EXIOBASE 3: Developing a Time Series of Detailed Environmentally Extended Multi-Regional Input-Output Tables, Journal of Industrial Ecology, 22(3). Retrieved from https://onlinelibrary.wiley.com/doi/10.1111/jiec.12715

Turubanova et al. (2018): Ongoing primary forest loss in Brazil, Democratic Republic of the Congo, and Indonesia, Environmental Research Letters, 13(7), 074028. Retrieved from https://doi.org/10.1088/1748-9326/aacd1c

WWF (2021): Assessing portfolio impacts. Retrieved from https://wwfint.awsassets.panda.org/downloads/wwf_assessing_portfolio_impacts_final.pdf

Warmerdam (2023): Forests & Finance financial research methodology. Retrieved from https://forestsandfinance.org/wp-content/uploads/2020/08/Forests-Finance-financial-research-methodology-01Sep2020.pdf

Weber et al. (2017): ASSET-LEVEL DATA AND CLIMATE-RELATED FINANCIAL ANALYSIS: A MARKET SURVEY. Retrieved from https://2degrees-investing.org/wp-content/uploads/2017/02/Asset-Level-Data-Market-Report.pdf

World Bank and WWF (2020): Spatial Finance Challenges and Opportunities in a changing world. Retrieved from https://documents1.worldbank.org/curated/en/850821606884753194/pdf/Spatial-Finance-Challenges-and-Opportunities-in-a-Changing-World.pdf

PUBLICATION DETAILS

CREATED BY

Climate & Company
Ahornallee 2 | 12623 Berlin
www.climateandcompany.org
hello@climcom.org
EU Transparency Register: 287394243384-58

Version 1.0

AUTHORS

Lead Authors:
Malte Hessenius (malte@climcom.org)
Marc Bohnet (marc@climcom.org)
Tycho Tax (tycho@climcom.org)

Contributing Authors:
Åsa Mossberg (AP2)
Fiona Pedeboy (Global Canopy)
Helen Bellfield (Global Canopy)
Pei Chi Wong (Global Canopy)

ACKNOWLEDGEMENTS

Climate & Company's work is funded by the Gordon and Betty Moore foundation. The Moore Foundation is a philanthropic organisation founded in 2000 by Intel co-founder Gordon Moore and his wife Betty. The Moore foundation is the world's biggest funder of nature-related work and has funded a variety of projects in the nature finance space such as Trase or Global Forest Watch Pro.

In this project Climate & Company actively collaboration with the junior research group PECan from Humboldt Universität zu Berlin. Funded by the German Ministry of Education and Research (BMBF), PECan explores the socio-ecological interactions between climate, political and financial systems.

Further acknowledgements: We thank Ingmar Juergens and Dr. Oliver Herrmann (Climate & Company) for helpful general advice. We also thank Martina Kaplanová and Robert Norris for helpful discussions on the integration of geospatial data.