

# THE GLOBAL BIODIVERSITY SCORE

GBS Review: Input-Output modelling

July 2020 – Corrected version

## Content

Note to the reader.....	2
<b>1 Context .....</b>	<b>2</b>
<b>1.1 Objective and overview of this report.....</b>	<b>2</b>
<b>1.2 Basics on input-output analysis .....</b>	<b>3</b>
<b>A The input-output framework.....</b>	<b>4</b>
<b>B Multi-regional and environmentally extended input-output models .....</b>	<b>5</b>
<b>C The EXIOBASE model.....</b>	<b>7</b>
<b>2 Default assessments .....</b>	<b>7</b>
<b>2.1 Characterisation factors.....</b>	<b>7</b>
<b>A Overall approach .....</b>	<b>8</b>
<b>B D matrices.....</b>	<b>10</b>
<b>C M matrices .....</b>	<b>17</b>
<b>D D+M matrices .....</b>	<b>17</b>
<b>E Computing the static and dynamic impacts per scope .....</b>	<b>18</b>
<b>2.2 Using activity data to assess impacts .....</b>	<b>20</b>
<b>A Dimensioning the impact .....</b>	<b>20</b>
<b>B Attributing.....</b>	<b>28</b>
<b>3 Refined assessments .....</b>	<b>34</b>
<b>3.1 Using company data on purchases.....</b>	<b>34</b>
<b>3.2 Integrating refined assessments from inventories, pressures, direct measurements ..</b>	<b>35</b>
<b>4 Limits and perspectives .....</b>	<b>36</b>
<b>4.1 Limitations of EXIOBASE 3 .....</b>	<b>36</b>
<b>4.2 Limitations regarding the use of IO modelling in the GBS and perspectives for future developments.....</b>	<b>37</b>
<b>References .....</b>	<b>39</b>

# Note to the reader

GBS review reports are not completely independent from each other. Readers of this report are advised to first read the report dedicated to **Core concepts of the GBS** (CDC Biodiversité 2020a) to ensure a good overall comprehension of the tool and the present report.

The following colour code is used in the report to highlight:

- Assumptions

- Important sections

- Developments of the GBS planned in the future

The GBS review reports are aimed at technical experts looking for an in-depth understanding of the tool and contribute to the transparency that CDC Biodiversité considers key in the development of such a tool. They focus on technical assumptions and principles. Readers looking for a short and easy-to-understand explanation of the GBS or on an overview of existing metrics and tools should instead read the general audience reports published by CDC Biodiversité (CDC Biodiversité 2017; CDC Biodiversité, ASN Bank, and ACTIAM 2018; CDC Biodiversité 2019b).

[ERRATUM]: Following the review, some examples and figures have been updated to reflect the modifications in the tool that had been made during the review. Some slight corrections have also been made, ie. double counting in the impacts of Food2 and Mining2 in the test portfolio

## 1 Context

### 1.1 Objective and overview of this report

This report presents the Input-Output (IO) modelling procedure used in GBS. The use of IO approach in the tool is two-fold

1. The IO framework allows to run default corporate and portfolio assessments based on limited activity data, namely the turnover by industry and country;

2. The IO model allows to integrate the value-chain impacts of activities based either on default or refined data (only upstream impacts are assessed yet).

The report is organized as follows: the link between the main sections of the report and the GBS framework is illustrated on Figure 1. Section 1.2 presents the basic IO framework and the IO model EXIOBASE used in the GBS. Section 2 describes how default assessments are undertaken, notably distinguishing the computation of characterisation factors (2.1) and the computation of default impacts based on activity data (2.2). Section 3 describes how default assessments can be improved based on refined company data whether monetary (3.1) or related to inventories and pressures (3.2).

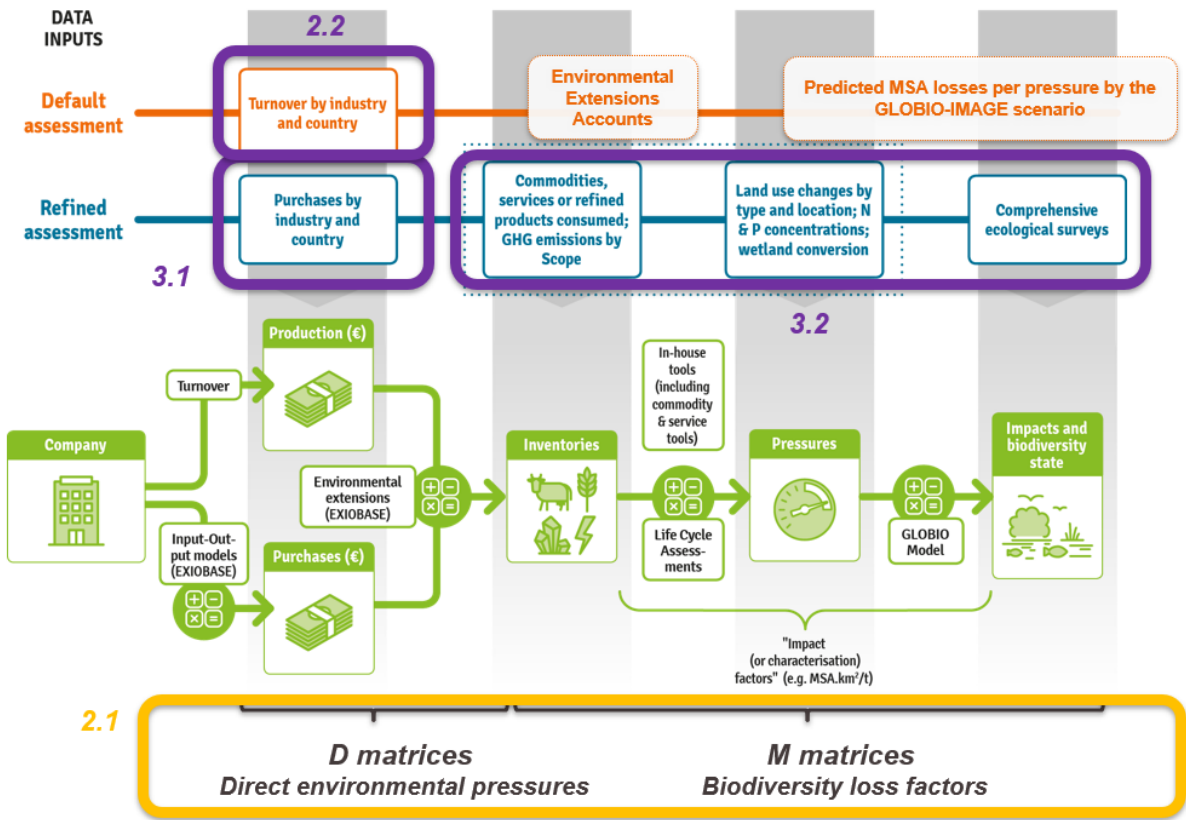


Figure 1: Link between the content of this report and the GBS framework

## 1.2 Basics on input-output analysis

## A THE INPUT-OUTPUT FRAMEWORK

Input-output (IO) models are constructed from observed economic data and provide information about the activity of industries that both produce and consume goods. Those interindustry relationships are derived from interindustry transaction tables in which the rows describe the composition of inputs required by a particular industry to produce its output. The data often concern industries in a specific geographic region (nation, state, county, etc.) and consists in monetary flows. Additional columns represent final demand, *i.e.* the sales to final markets (consumers, government, exports), and additional lines deal with value added, *i.e.* account for non-industrial production inputs (labor, capital, etc.)<sup>1</sup>.

An IO model is a system of linear equations where the total output of an industry is distributed through sales to other sectors and to final demand, that is

$$x_i = z_{i1} + \dots + z_{ij} + \dots + z_{in} + f_i = \sum_{j=1}^n z_{ij} + f_i,$$

where

- $x_i$  is the monetary value of the total output of sector  $i$  during the year;
- $z_{ij}$  is the demand of sector  $j$  for inputs from sector  $i$  during the year;
- $n$  is the number of industries in the economy;
- $f_i$  is the total final demand for sector  $i$ 's production during the year.

Such an equation exists for the  $n$  industries of the economy, so that the set of  $n$  linear equations with  $n$  unknowns can be represented using the matrix representation

$$\mathbf{x} = \mathbf{Z}\mathbf{i} + \mathbf{f}$$

with  $\mathbf{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$ ,  $\mathbf{Z} = \begin{bmatrix} z_{11} & \dots & z_{1n} \\ \vdots & \ddots & \vdots \\ z_{n1} & \dots & z_{nn} \end{bmatrix}$ ,  $\mathbf{f} = \begin{bmatrix} f_1 \\ \vdots \\ f_n \end{bmatrix}$  and  $\mathbf{i}$  is the summation vector (column vector of ones with size  $n$ ).

The interindustry flows provide information on the amount of each input required to produce the output, *i.e.* the production functions. The ratio of input required over the total output (in \$/\$) is called the technical coefficient and is computed as

$$a_{ij} = \frac{z_{ij}}{x_j} = \frac{\text{value of input } i \text{ bought by sector } j}{\text{value of total production of sector } j}.$$

The linear equations can thus be re-written as

$$x_i = a_{i1}x_1 + \dots + a_{ij}x_j + \dots + a_{in}x_n + f_i = \sum_{j=1}^n a_{ij}x_j + f_i.$$

---

<sup>1</sup> Within a country, the sum of the value added is equal to the Gross Domestic Product (GDP).



Denoting  $\mathbf{A}$  the matrix of technical coefficients we thus have

$$\mathbf{x} = \mathbf{Ax} + \mathbf{f},$$

$$\mathbf{x} = (\mathbf{I} - \mathbf{A})^{-1}\mathbf{f}.$$

The matrix  $(\mathbf{I} - \mathbf{A})^{-1}$  is called the Leontief inverse matrix. Notice that the Leontief inverse matrix can be seen as a power series approximation

$$(\mathbf{I} - \mathbf{A})^{-1} = \mathbf{I} + \mathbf{A} + \mathbf{A}^2 + \mathbf{A}^3 + \dots + \mathbf{A}^n + \dots = \lim_{n \rightarrow +\infty} \sum_{k=0}^n \mathbf{A}^k.$$

The Leontief inverse matrix thus represents all the value chain interrelations required to produce the output.

## B MULTI-REGIONAL AND ENVIRONMENTALLY EXTENDED INPUT-OUTPUT MODELS

In multi-regional input-output (MRIO) models, data are spatialised and flows between geographical regions are detailed. Hence MRIO models prove useful to analyse global supply-chains. They are particularly suited to the analysis of global supply-chain-related environmental pressures. Indeed, it has been extended to account for pollution generation and abatement associated to economic activity since the late 1960s. Environmental – and social – “extensions” allow the comprehensive examination of a wide variety of factors – employment, pollution, water, capital expenditures, etc. – associated to economic activities and policies. Environmental extension of MRIO models can be done in several ways, either by augmenting the technical coefficient matrix with additional rows and/or columns reflecting environmental impacts or by including additional “ecosystem” sectors where flows between these sectors and economic sectors are recorded along the lines. Whichever method is used, EEMRIO models often mix physical and monetary units as pollutants emissions and volumes of raw materials sourced are linked to the monetary value of production.

Environmentally extended multi-regional input-output (EEMRIO) models are used today to compute climate, water and material resources footprints of production and consumption at national or regional levels. In short, while MRIO models provide a mathematical representation of the flows of goods and services between industries all over the world by documenting the monetary transactions involved in production and consumption, EEMRIO models are appropriate for analysing the supply-chain-related environmental pressures due to production and consumption activities. Indeed in EEMRIO models, companies’ purchases, commodity extraction and emissions are spatially explicit, so that complex international value chains can be analysed.

EEMRIO models provide data on material, water and land-use consumptions and emissions of substances related to the economic activities of a detailed list of industries all over the world (the “Environmental extensions” in Figure 2). They are recognised as key frameworks to provide a comprehensive description of the global economy and analyse its effect on the environment and are thus interesting tools to support biodiversity footprint methodologies. In fact, (Wilting and van Oorschot 2017a; Wilting et al. 2017) use such a framework to quantify biodiversity losses due to, respectively, consumption and production activities in

the Netherlands. EEMRIO models display several advantages in the development of the GBS compared to, for instance, the life-cycle analysis framework:

1. They allow the evaluation of companies based solely on the distribution of their turnover across region and industry;
2. Conceptually, all industries of all countries can be evaluated simultaneously;
3. They are self-contained and guarantee internal consistency between and within monetary as well as physical amounts.

Since these models are based on a sectoral approach, they face limits when analysing companies within the same industry. Indeed, in an EEMRIO framework, two companies operating in the same industries and the same regions cannot be distinguished otherwise than by the monetary value of their production. Thus, three main applications of EEMRIO tools in the GBS are envisaged:

- Providing **benchmark industry footprints at the national level**;
- Calculating a **generic corporate footprint including value chain impacts** for companies with limited information. The footprint obtained would be considered as a “default footprint” with room for improvement if the company chooses to disclose more specific data. As explained below, computing the Scope 1, Scope 2 and Scope 3 footprint of a company using EEMRIO data only requires the knowledge of the breakdown of its activity by country and industry. The main industry and the location of the headquarter of each company is usually known. The distribution of turnover by country and industry can sometimes be deduced from the company’s annual reports but is not always available;
- Calculating the **footprint of financial assets**. The footprint of the companies financed is first assessed (“dimensioning” of the impact), then attribution rules specific to the asset class are applied to evaluate the impact which can be attributed to the funding source (“attribution” of the impact). The monetary framework of EEMRIO proves especially appropriate in this context.

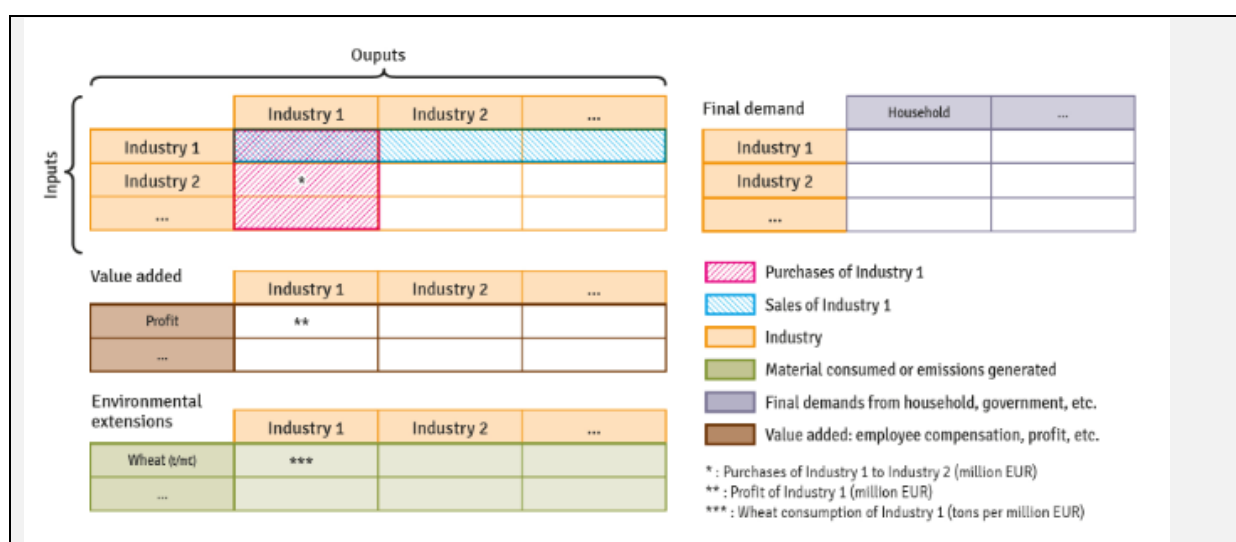


Figure 2: Components of an environmentally extended multi-regional input-output model

## C THE EXIOBASE MODEL

The EEMRIO model used in the GBS is EXIOBASE version 3.4 (Stadler et al. 2018), noted “EXIOBASE 3” in the rest of the report. It is a time series of EEMRIO tables ranging from 1995 to 2011 for 49 regions (44 countries and 5 rest of the world regions) and 163 industries. The list of regions and industries is reproduced in (CDC Biodiversité 2019a). Though the times series provide information for analysing the dynamics of environmental pressures of economic activities over time, only the data for year 2011 are used in the GBS. Data related to environmental impacts are grouped into 4 accounts:

- The **emission account** provides quantitative data on industry-specific emissions of 27 pollutants, including GHGs, nitrogen and phosphate;

- The **water account** documents water consumption (blue and green) and water withdrawal of agricultural, manufacturing and energy production activities;

- The **material account** documents the extraction of 222 raw materials, including biomass items (data retrieved from FAOSTAT), metal ores and minerals (data retrieved from the British Geological Survey (BGS 2014), the US Geological Survey (USGS 2014) and the World Mining Data (Reichl, Schatz, and Zsak 2014)) and fossil fuels (data retrieved from the International Energy Agency (IEA 2014a; 2014b));

- The **land account** lists the area consumption related to agricultural and settlement activities for 15 types of land use.

The data can be downloaded for free and are available in two formats. Industry by industry tables describe interindustry monetary flows, while product by product tables describe the relationships between products directly. Since we are interested by the sectoral approach, we downloaded the industry by industry IO tables for the year 2011. The folder contains (main files):

- Y.txt: final demand, *i.e.* sales of the 163 industries of the 49 regions to the 7 types of final demand in each region
- A.txt: the matrix of technical coefficients
- F.txt: the environmental accounts for all extensions (energy, emissions, water, materials, land) related to the total production (and not the production per million euros) of each industry
- unit.txt: file describing the units of all figures in the data

## 2 Default assessments

### 2.1 Characterisation factors



## A OVERALL APPROACH

Conceptually we define

$$impact_j = i(M \circ D)(I - A)^{-1}x_j,$$

where

- $impact_j$  is the biodiversity footprint of sector  $j$  in MSA.km<sup>2</sup>
- $i$  is a  $(1 \times t)$  vector of ones used to sum up the biodiversity losses caused by individual environmental pressure
  - $t$  is the number of pressures considered
- $M$  is the  $(t \times r.s)$  **matrix of biodiversity loss factors**,  $m^{i,j}$  is a  $(1 \times s)$  row vector of biodiversity loss factors of direct environmental pressure  $i$  in region  $j$  (depicting the biodiversity losses per unit of environmental pressure)
  - $t$  is the number of pressures considered
  - $r$  is the number of regions in the IO model
  - $s$  is the number of industries in the IO model
- $D$  is the  $(t \times r.s)$  **matrix of direct environmental pressures**,  $d^{i,j}$  is a  $(1 \times s)$  row vector of direct environmental pressure intensities of pressure  $i$  in region  $j$  (depicting the direct environmental pressures of one unit of production for all sectors)
- $(I - A)^{-1}$  is the Leontief inverse matrix (cf above)
- $x_j$  is the  $(r.s \times 1)$  vector of production with zeros for all sectors except sector  $j$
- Operation  $\circ$  is the element-wise multiplication of two matrices (Hadamard product)

The main components of the equation are the M and D matrices and the Leontief inverse. The Leontief inverse is computed based on the matrix of technical coefficients, as explained in Section 1.2A. Since the L matrix is not part of the files provided for download on EXIOBASE website, the researchers of the Dutch Institute of Environmental Sciences (CML, EXIOBASE developers) kindly shared the MATLAB code required to compute it based on the A matrix (file A.txt). The main lines of the code are reproduced below.

MATLAB CODE TO COMPUTE LEONTIEF INVERSE MATRIX – SOURCE: ARJAN DE KONING, CML

```
A = dlmread('A.txt','\t',3,2);
disp('Ready reading "mrIot.txt"');
[rows, cols] = size(A)
I = eye(cols);
Linv = inv(I-A);
disp('Ready calculating Leontief inverse');
```

The computation of D, M and D◦M matrices is detailed in the other subsections. Before coming to them, we describe the overall IO approach based on the theoretical concepts introduced above.

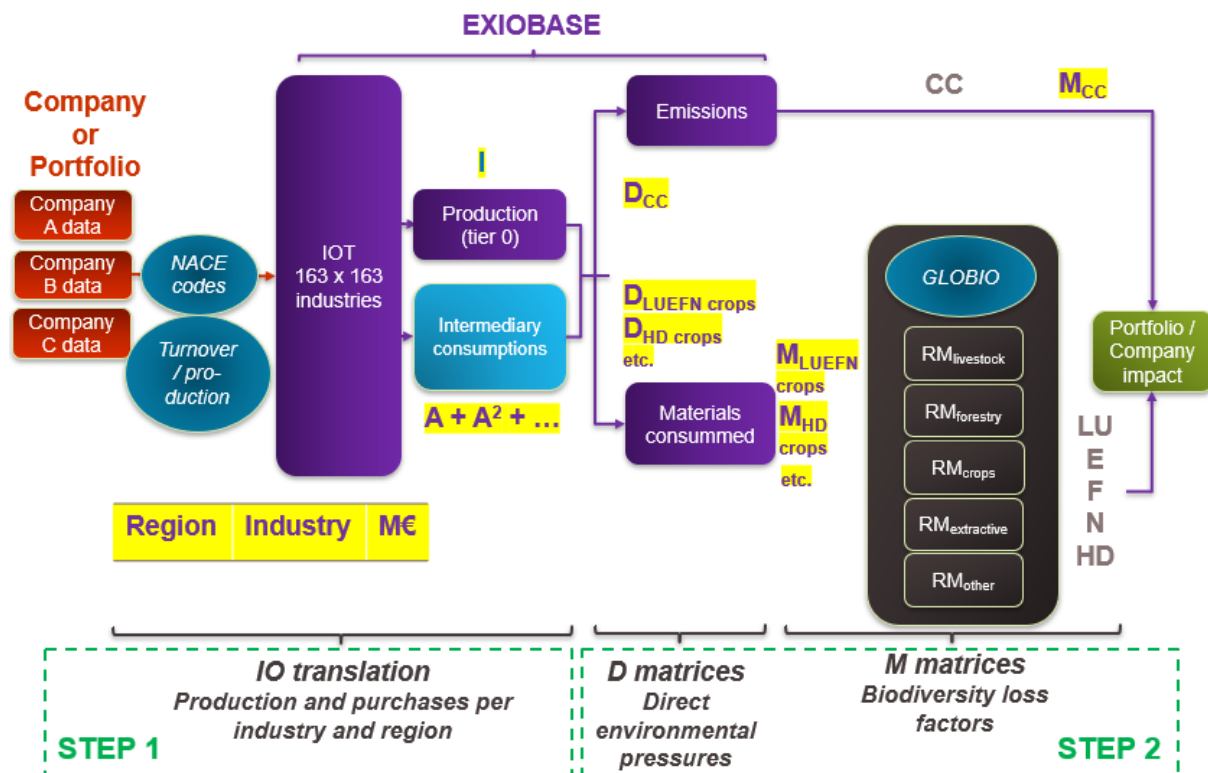


Figure 3: Detailed GBS IO approach

The overall approach is illustrated by Figure 3, including the required data and use of the various components of EXIOBASE IO model and the D and M matrices introduced above. The **first step** of the analysis is the combination of company or portfolio data with the MRIO model EXIOBASE 3 to translate them into production and purchases per industry and region, in order to assess the impact of the whole upstream value chain. The **second step** of the default assessment is to link the production of EUR 1 million of any industry and any region to direct (Scope 1) biodiversity impacts. This analysis can be broken down in two sub-steps:

### 1. The assessment of the inventory data of the production of EUR 1 million worth of output of any industry in any region.

➤ This component gives information on the contributions of the activity to drivers of biodiversity loss, mostly the emission of GHGs and the consumption of raw materials and water. It is calculated based on the environmental extensions of the EEMRIO model EXIOBASE 3. These “direct environmental pressures” include GHG emissions, water and material consumptions, among others. They are gathered in what we call the “D matrices”, following Wilting & van Oorschoot (2017). In the life-cycle analysis world, they would be called “inventory data”, as in Figure 1

### 2. The assessment of the “biodiversity impacts” of the drivers.

➤ This component gives information on the loss of biodiversity caused per unit of driver (kg CO<sub>2</sub>-eq, ton of raw material) in MSA.km<sup>2</sup>. We call it the “M matrices”, following Wilting & van Oorschot (2017). Biodiversity impact factors spatially explicit and calculated based on the combination of the GLOBIO model and several commodity-specific tools developed within the GBS (e.g. the “Crops” CommoTool).

The assessment of environmental pressures and biodiversity impacts are conducted simultaneously for all industries and regions thanks to the dedicated D and M matrices.

The **result** is the Scope 1 and upstream supply chain biodiversity impacts of the company or portfolio studied. Downstream impacts, notably impacts related to final demand (e.g. consumers, public administrations) are not assessed yet.

The default assessment can be used at various levels of the GBS methodology, either for default corporate assessments or in refined assessments when a company directly provides financial data about its purchases (see Figure 1. The IO framework also enables the distinction of the various tiers<sup>2</sup> of the supply chain (direct suppliers, suppliers higher in the value chain).

## B D MATRICES

D matrices depict the contribution of economic activities to the pressures on biodiversity, e.g. GHG emissions and raw material consumption per million euros of output. They document **direct environmental pressure factors** linking the production of EUR 1 million worth of output of each {region; industry} to the related emissions (GHGs, pollutants, etc.) and extractions (raw materials, water, etc.). These factors are thus **quality tier 2 factors for raw materials** and **data quality tier 1 for GHGs**. EXIOBASE data provide the total amounts for each account (emissions, materials, etc.) per {region; industry}, e.g. the total CO<sub>2</sub> emissions due to the production of the French industry “Pigs farming” for the year 2011. A prerequisite for creating the D matrices is thus the computation of the total production of each industry in each region. The total amounts of emissions and materials are then divided by the total monetary production per {region; industry} to get the amount per million euros.

Following the IO approach, the total production of each {region; industry} can be computed based on the vector of total final demand and the Leontief inverse as:  $\mathbf{x} = \mathbf{L} \cdot \mathbf{f}$  with the following R code:

### COMPUTING TOTAL PRODUCTION

```
### Load data and keep only the numeric data from final demand
Y <- read.table("Path\Y.txt", header = FALSE, sep = "\t", stringsAsFactors = FALSE)
final_demand <- Y[-c(1:3), -c(1,2)]
### Compute total production of each industry per region
# 1) Convert data into matrices
```

<sup>2</sup> “Tier” here refers to supplier tier (tier 1 are direct suppliers, tier 2 are the direct suppliers of a company’s tier 1 suppliers, etc.). This is a totally different concept from the “data quality tiers” introduced in the Introduction document (CDC Biodiversité 2020a).

```

260 # [...]
261 # 2) Compute the vector of total final demand
262 # row vector with total final demand (all categories) per region and industry
263 total_fd <- as.vector(apply(t(final_demand_matrix), 2, sum))
264
265 # 3) x = L.f
266 # L_inverse_matrix_integer = L * 10^6 to reduce the size of the database so need to
267 divide
268 by 10^6
269 total_production <- ((GBStoolbox::L_inverse_matrix_integer) / 10^6) %%% total_fd

```

### 2.1.B.1 Climate change

Because 1) GHG emission data are computed directly per million euros based on EXIOBASE emissions account and 2) an impact factor in MSA.km<sup>2</sup>/kgCO<sub>2</sub>-eq can be computed based on GLOBIO data (see Section 0), the impact due to climate change is computed using directly the pressure-impact relationship. The same kind of reasoning is true also for hydrological disturbance since the impact factors are expressed per million of cubic meters withdrawn or consumed (see Section 2.1.C.3 and (CDC Biodiversité 2020c)). On the contrary, for the other terrestrial and aquatic pressures the impact factors are computed per ton of material extracted in the CommoTools. Consequently, the D matrices related to these pressures should document tons of materials per million euros of production.

The D matrix related to climate change (noted D<sub>cc</sub>) provides the amount of GHG emissions related to the production of EUR 1 million of each {region; industry} in kg CO<sub>2</sub>-eq/mEUR. It is computed based on the emission account which documents the total emissions of several substances of each industry. The computation steps are

1. Compute the emission account per million euros for each substance
2. Select the GHGs among the substances listed
3. Convert the quantities emitted into kg CO<sub>2</sub>-eq using their global warming potentials
4. Sum the emissions over all gases per {region; industry}

We consider emissions of the six gases covered by the Kyoto Protocol, *i.e.* carbon dioxide (CO<sub>2</sub>), fossil and biogenic methane (CH<sub>4</sub>), nitrous oxide (N<sub>2</sub>O), sulphur hexafluoride (SF<sub>6</sub>), hydrofluorocarbons (HFCs) and perfluorocarbons (PFCs). GHG emissions are expressed in CO<sub>2</sub>-equivalents using Global Warming Potentials (GWPs) taken from the GHG Protocol (Table 1). For the GWP, we consider a time horizon of 100 years in the calculations, consistent with the IPCC (Stocker 2014). Since emissions of HFCs and PFCs are already given in kg CO<sub>2</sub>-eq in EXIOBASE, no GWP is used.

Table 1: Global Warming Potential of the main GHGs for a time horizon of 100 years, source: (Stocker 2014)

Greenhouse gas	GWP (kg CO <sub>2</sub> -eq/kg) for 100 years
CO <sub>2</sub>	1
CH <sub>4</sub>	28
N <sub>2</sub> O	265
SF <sub>6</sub>	23 500

295 EXIOBASE environmental extension distinguishes emissions due to combustion from emissions related to  
 296 non-combustion sources. Both types of emissions are kept separated in the D matrix. Indeed, both types of  
 297 emissions are used to compute the climate related impact, but only combustion related emissions are  
 298 considered to compute that of Scope 2 (see Section 2.1.E).

```

299 COMPUTING D_CC – MAIN STEPS
300 # 1/ Get GHG emissions and convert GHG emissions into kg CO2 eq.
301 # Step 1: EE_data_emissions: environmental extension emissions account
302 GHG_emissions <- EE_data_emissions %>%
303   # Step 2: keep only lines concerning GHGs
304   filter(str_detect(Item,"CO2|CH4|N2O|SF6|HFC|PFC")) %>%
305   # extract the name of the GHG (3 first characters in emission name)
306   mutate(GHG = substr(Item, 1, 3)) %>%
307   right_join(GHG, by = "GHG") %>% # merge with GHG database
308   select(-GHG_Name, -Comment, -GHG_category) %>% # drop unwanted columns
309   # Step 3: convert into kgCO2eq
310   mutate(Emissions_kg_CO2_eq = Total_quantity * Coefficient_100) %>% #
311   select(-Coefficient_100) %>% # drop Coefficient variable
312   arrange(ID_region, ID_industry) # sort by region and industry
313
314 # 2/ Distinguish combustion and non-combustion
315 # Extract combustion emissions
316 combustion_emissions <- GHG_emissions %>% filter(str_detect(Item, "- combustion -"))
317 # Extract non-combustion emissions, which include both emissions flagged as "non
318 combustion" but also emissions with no "combustion" or "non combustion" flags, e.g. "CH4 -
319 waste - air"
320 non_combustion_emissions <- GHG_emissions %>% filter(!str_detect(Item, "- combustion -"))
321
322 # 3/ Step 4: Sum by industry, region and GHG and join with total_production to get the
323 production in M€ and calculate the emissions per M€
324 # COMBUSTION
325 combustion_emissions <- combustion_emissions %>%
326   group_by(ID_region, ID_industry, GHG) %>%
327   summarise(Emissions_kg_CO2_eq = sum(Emissions_kg_CO2_eq)) %>%
328   # Compute the emissions per M.EUR using total production
329   left_join(total_production, by=c("ID_region", "ID_industry")) %>%
330   mutate(Combustion_emissions_per_mEUR = Emissions_kg_CO2_eq / Total_production)
331 # NON-COMBUSTION
332 # same as above on "non_combustion_emissions"
333
334 # 4/ Join the 2 sub-tables to have 1 table with combustion and non-combustion emissions
335 and sum by industry and region
336 D_CC <- combustion_emissions %>%
337   full_join(non_combustion_emissions, by=c("ID_region", "ID_industry", "GHG")) %>%
338   arrange(ID_region, ID_industry, GHG) %>%
339   # Total emissions per mEUR per (industry; region)
340   group_by(ID_region, ID_industry) %>%
341   summarise(Combustion_emissions_per_mEUR = sum(Combustion_emissions_per_mEUR,
342     na.rm = TRUE),
343     Non_combustion_emissions_per_mEUR = sum(Non_combustion_emissions_per_mEUR,
344     na.rm = TRUE))
345

```

## 2.1.B.2 Land use, encroachment, fragmentation, nitrogen deposition

The D matrix related to land use (LU), encroachment (E), fragmentation (F) and nitrogen deposition (N), noted  $D_{LUEFN}$ , provides the amount of each raw materials<sup>3</sup> related to the production of one million euros of each {region; industry} in t/mEUR. It is computed based on the materials account which documents the total quantity of raw materials extraction for each industry. Raw materials can be split in broad categories corresponding to the CommoTools developed for the GBS: primary crops, non-primary crops (fodder, crop residues), grazing, wood logs, metal ores, non-metallic minerals and fossil fuels. Some categories require more computation steps than other, as we will explain below. Hence the distinction between  $D_{LUEFN}$  matrix, a broad mother-matrix that results from the computation steps common to all raw material categories, and the subsequent  $D_{LUEFN\_category}$  matrices that are the end-results used in the assessments.

### 2.1.B.2.1 Computation of $D_{LUEFN}$ mother matrix

The computation of  $D_{LUEFN}$  is straightforward. Starting from the environmental extension related to raw materials (converted from kilotonnes into tonnes), the total quantity of each raw material extracted is simply divided by the total output per {region; industry} to get the extraction in t/mEUR. One subtle point is that EXIOBASE distinguishes two types of extractions: the “used domestic extraction” and the “unused domestic extraction”. The first type documents the actual quantity of raw material extracted (e.g. tonnes of carrots) while the second one refers to the amounts that come with the raw materials of interest but are not used (e.g. stems, leaves and other residuals of biomass extraction, overburden for mining activities). The biodiversity impact factors in the M matrices, computed based on CommoTools outputs, consider tons of used materials: tons of actually cultivated carrots, tons of extracted ores (without the overburden which is estimated separately). Therefore, only the domestic extraction is considered in the  $D_{LUEFN}$  matrices to ensure consistency.

```
COMPUTING D_LUEFN
D_LUEFN <- EE_data_materials %>% # EXIOBASE materials account
  # keep only the used extraction because "unused extraction" does not correspond to
  # additional raw materials but rather to extractions that we don't want to consider
  # (overburden in mining, leaves and stems for crops...)
  filter(str_detect(Item, "Domestic Extraction Used")) %>%
  left_join(total_production, by = c("ID_region", "ID_industry")) %>%
  # compute the quantity of raw material extracted per mEUR
  mutate(Quantity = Total_quantity / Total_production,
         Item = str_replace(Item, "Domestic Extraction Used - ", "")) %>%
  select(ID_region, ID_industry, ID_region_group, ID_industry_group, Item, Quantity)
```

<sup>3</sup> Designated by the variable “Item” in the code. The raw format of Item is “Type of extraction – Raw material category – Raw material”, e.g. “Domestic Extraction Used – Primary Crops – Avocados”.



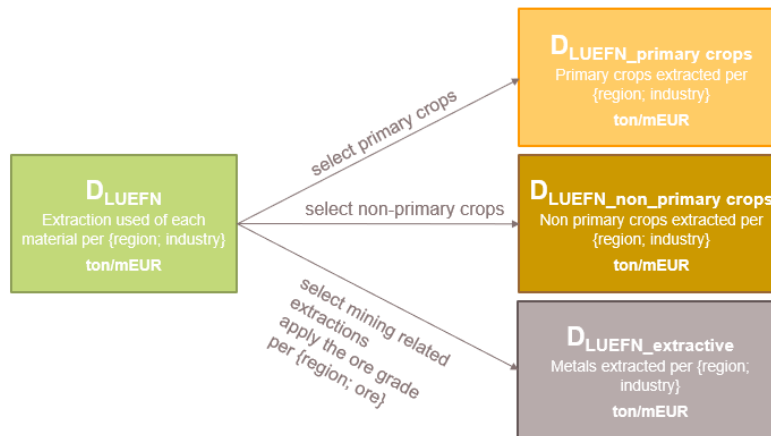


Figure 4: Link between  $D_{LUEFN}$  mother matrix and the matrices used in GBS

Figure 4 illustrates how final D matrices are computed based on  $D_{LUEFN}$  mother matrix.

#### 2.1.B.2.2 Computation of $D_{LUEFN\_primary\_crops}$ and $D_{LUEFN\_non\_primary\_crops}$

$D_{LUEFN\_primary\_crops}$  and  $D_{LUEFN\_non\_primary\_crops}$  are computed simply by filtering  $D_{LUEFN}$  to keep the items of the desired category<sup>4</sup>.

#### 2.1.B.2.3 Computation of $D_{LUEFN\_extractive}$

On the contrary  $D_{LUEFN\_extractive}$ , the D matrix related to metal ores, requires more work. Indeed, the “used” extraction provided by EXIOBASE is the total quantity of extracted gross ore without the overburden (cf. above) but including gangue, not the quantity of the ore of interest. The quantity of extracted gross ore is computed by EXIOBASE team based on production data per metal taken from the British Geological Survey (BGS 2014) and an estimation of ore densities obtained through interviews with experts and a literature review. Since the GBS CommoTool gives the impact per tonne of metal (CDC Biodiversité 2020d), we need to correct the data so that  $D_{LUEFN\_extractive}$  documents tonnes of metal. For now, we do so by working EXIOBASE computation backwards:

1. Get BGS data and compute the total production of each metal per EXIOBASE region
2. Compute the ore density per {region; metal} by dividing the total production by the gross ore extraction
3. Apply the computed ore grades to the gross ore extraction data in  $D_{LUEFN}$  to get the corresponding metal extraction

A path to improve the methodology in the future would be to be able to use directly the ore grades used by EXIOBASE instead of re-computing them.

<sup>4</sup> This is fairly easy since the raw material category is documented in the Item name. In the end matrices, only the raw material name is kept from the original name.

The 9 metals considered in the CommoTool are aluminium, copper, iron, gold, lead, nickel, silver, tin and zinc. They correspond to separate mining industries in EXIOBASE and their extraction is separately documented in the environmental extensions, except for aluminium which corresponding industry is “Mining of bauxite and aluminium” and raw material is “Bauxite and aluminium”. We thus need to make an assumption on the proportion of bauxite and aluminium in the raw material extracted.

#### ASSUMPTION

The proportion of bauxite in the raw material “Bauxite and aluminium” is 100%. The corresponding quantity of aluminium is computed using the ratio of aluminium to bauxite, which is 16.5%.

The main code lines involved in the computation of regional ore grades and  $D_{LUEFN\_extractive}$  are reproduced below. To ensure data consistency, the computed grades are compared with the highest known grade for each metal<sup>5</sup>. When the computed grades are higher than 1.5 the highest known grade<sup>6</sup>, they are replaced by the maximum between the world average and half the highest grade. This replacement procedure allows to 1) maintain differentiation between mines (world average is not the only replacement figure and a high grade is allocated to mines for which a high grade was computed), 2) ensure that abnormal grades are controlled for (computed grades higher than 1.5 the highest grade are replaced), and 3) stick to a rather conservative approach (for mines with a very high computed grade, only half the world's highest grade is used rather than the highest grade). Other abnormal computed grades (superior to 1 for instance) are replaced by the average grade.

#### COMPUTE ORE GRADE PER EXIOBASE REGION

```
ore_grade_per_exiobase_region <- bgs_2011_production %>%
  # Link each BGS country to the corresponding EXIOBASE region
  # [...]
  # compute the ore grade per region and metal based on the production documented on BGS
  # data and the extraction documented in EXIOBASE materials account
  group_by(ID_region, commo_name) %>%
  mutate(region_production = sum(Production_tons, na.rm = TRUE),
         ore_grade = region_production / extracted_tons) %>%
  # [...]
  # convert bauxite into aluminium
  mutate(ore_grade = case_when(
    commo_name == "Bauxite" ~ ore_grade * bauxite_to_alu_ratio,
    TRUE ~ ore_grade),
    commo_name = if_else(commo_name == "Bauxite", "Aluminum", commo_name)) %>%
  # computed grade analysis based on ore characteristics
  left_join(extractive_ore_specs, by = "commo_name") %>%
  group_by(commo_name) %>%
```

<sup>5</sup> The highest ore grades for each metal can be found online rather easily on specialized sites like mining.com which provide rankings of highest-grade mines for several ores (copper, gold, lead, silver, zinc) based on private data from Mining Intelligence. We compare to 1.5 x highest grade to allow for uncertainty around the highest grade.

<sup>6</sup> For iron we compare to the highest grade instead of 1.5 x highest grade because the highest known grade is 0.7, hence 1.5 x highest grade would be superior to 1.

```

439 # when the computed grade is higher than 1.5*highest_grade, we replace by the max between
440 highest_grade/2 and average_grade
441 mutate(ore_grade = case_when(
442   ore_grade > 1.5 * commo_grade_highest & commo_name != "Iron" ~
443     max(commo_grade_highest / 2, commo_grade),
444   ore_grade > commo_grade_highest & commo_name == "Iron" ~
445     max(commo_grade_highest / 2, commo_grade),
446   # use average values when the computed grade is obviously weird
447   # [...]
448
449 COMPUTE D_LUEFN_EXTRACTIVE
450 D_LUEFN_extractive <- D_LUEFN %>%
451   # keep only the 9metals considered in the commotool
452   # [...]
453   # compute the amount of metal extracted
454   left_join(ore_grade_per_exiobase_region, by = c("ID_region", "commo_name")) %>%
455   mutate(Quantity = Quantity * ore_grade)

```

#### 2.1.B.2.4 Computation of D<sub>LUEFN\_wood\_logs</sub>

Text will be added here in future reports

#### 2.1.B.2.5 Computation of D<sub>LUEFN\_oil\_and\_gas</sub>

Text will be added here in future reports

### 2.1.B.3 Hydrological disturbance

The impact factors related to hydrological disturbance due to water abstraction (consumption and withdrawal) are expressed in MSA.km<sup>2</sup> per cubic meter of water withdrawn and consumed (CDC Biodiversité 2019c). The D matrix related to this pressure, noted D<sub>HD</sub> should thus document the cubic meters of water withdrawn and consumed per million euros of production. Its computation based on EXIOBASE environmental extensions is rather straightforward since the water account documents, among others, the total quantity of water consumed and withdrawn per {region; industry} in millions of cubic meters. As for GHG emissions, dividing by the total monetary production of the industry gives the cubic meters consumed and withdrawn per million euros. A subtlety however is that the water account distinguishes extractions of green water (coming from evapotranspiration of the vegetation), blue water (coming from water bodies) and grey water (post-industrial processes) for each {region; industry}. Considering that the impact factors concern the water abstracted from rivers and wetlands, only blue water is considered.

## C M MATRICES

As defined in Section 2.1.A, M matrices contain the impact factors per unit of environmental pressure. In GBS framework, M matrices contain the biodiversity impact of one unit of kg CO<sub>2</sub>-eq, one ton of any raw material and one million of cubic meters of water consumed or withdrawn. The impacts are expressed in MSA.km<sup>2</sup>/unit and are spatialized at the EXIOBASE region level. They thus fall into [the data quality tier 2](#).

The computation of M matrices is more straightforward than that of D matrices since it only requires to aggregate the results of the CommoTools at the desired geographical scale.

### 2.1.C.1 Climate change

The M matrix for climate change,  $M_{CC}$ , is simple since the impact of 1 kg CO<sub>2</sub>-eq is the same everywhere. Since the impact factor is  $4.37 \cdot 10^{-9}$  MSA.km<sup>2</sup>/kg CO<sub>2</sub>-eq (CDC Biodiversité 2020f),  $M_{CC}$  is thus simply a column vector of size  $(r \times s) = 49 \times 163 = 7987$  lines with the impact factor  $4.37 \cdot 10^{-9}$ .

### 2.1.C.2 Land use, encroachment, fragmentation, nitrogen deposition

The crop CommoTool provides impact factors for each FAO primary and non-primary crop in MSA.m<sup>2</sup>/ton at the country level (CDC Biodiversité 2020b). The items in EXIOBASE material accounts, and thus in  $DLUEFN_{crops}$  and  $DLUEFN_{non\_primary\_crops}$  are the same as that of the crop CommoTool since EXIOBASE uses FAOSTAT data. The item correspondence is thus straightforward. Computing  $MLUEFN_{crops}$  and  $MLUEFN_{non\_primary\_crops}$  therefore only requires the aggregation of the crop CommoTool impact factors from the country level to EXIOBASE region level. The correspondence between FAO countries and EXIOBASE regions is presented in (CDC Biodiversité 2019a). For each crop, the regional impact factor is computed as the weighted average of the impact factors of the countries in the region using the share of the country production in the total regional production of the crop as weight. A unit conversion is applied to express the impact factors in MSA.km<sup>2</sup>/ton.

The procedure is identical for the computation of biodiversity impact factors of metals.  $MLUEFN_{extractive}$  aggregates the results of the extractive CommoTool (CDC Biodiversité 2020d).

### 2.1.C.3 Hydrological disturbance

The impact factors of  $M_{HD}$  aggregate the results of the CommoTools for the pressure hydrological disturbance.

## D D◦M MATRICES

D◦M matrices ( $D\_x\_M$  in the code) are the Hadamard products of matching D and M matrices. Since D matrices transform economic activity into contribution to pressures on biodiversity and M matrices translate contributions into impact on biodiversity; combining both matrices translates million euros into biodiversity losses. The impact factors are therefore expressed for each {region; industry} in MSA.km<sup>2</sup>/mEUR. The

computation of D•M matrices is done simply by 1) multiplying the units of pressures (in D matrices, e.g. tons of each raw material consumed) by their impacts (in M matrices, e.g. the impact of the extraction of 1 ton of the raw material considered) and 2) aggregating the impacts at the {region; industry} level (e.g. summing for all the raw materials consumed by the {region; industry}).

## E COMPUTING THE STATIC AND DYNAMIC IMPACTS PER SCOPE

Impacts assessed through the direct environmental impacts and biodiversity impacts matrices are broken down between Scope 1, Scope 2 and Scope 3. The function `IO_evaluator_activity` computes Scope 1, Scope 2, Tier 1 of upstream Scope 3 (noted Tier 1 in subsequent text) and whole upstream Scope 3 static and dynamic impacts for all pressures. The results are computed by separate sub-functions and stored in separate data frames.

**Scope 1** impacts correspond to the impacts related to the company's production, *i.e.* its own turnover in the various {region; industry} pairs where it operates, not considering the related purchases. Computing the Scope 1 impact thus requires only to multiply the company's turnover (broken down by {region; industry}) by the D•M matrices. The typical code is presented below.

```
COMPUTE SCOPE 1 IMPACT (DYNAMIC OR STATIC)
# p replaces any pressure subscript. Amount is the turnover in the {region; industry}
scope1_impact <- pre_treated_activity_production %>%
  left_join(D_x_M_p, by = c("ID_region", "ID_industry")) %>%
  mutate(F_p = D_x_M_p * Amount) %>%
  # regroup lines that were split in the pre-treatment
  group_by(ID_line, ID_relationship, Name_business,
            ID_region, ID_region_group, ID_industry, ID_industry_group, Amount) %>%
  summarise(F_p = sum(F_p, na.rm = TRUE))
```

**Scope 2** impacts refer to the biodiversity impacts of the generation of the electricity, steam, heat and cold purchased. By definition, it includes only tier 1 suppliers, *i.e.* the ones selling directly to the company, and not those generating energy for the companies' other suppliers. These purchases in the various {region; industry} pairs are identified thanks to the IO tables which display the purchases to energy generation industries<sup>7</sup>. For now, only the climate change impacts of Scope 2 are considered, defined as the impacts due to combustion related GHG emissions (EXIOBASE distinguishes emissions due to fossil fuel combustion

<sup>7</sup> The energy production industries are: production of electricity by coal, production of electricity by gas, production of electricity by nuclear, production of electricity by hydro, production of electricity by wind, production of electricity by petroleum and other oil derivatives, production of electricity by biomass and waste, production of electricity by solar photovoltaic, production of electricity by solar thermal, production of electricity by tide, wave, ocean, production of electricity by geothermal, production of electricity nec, steam and hot water supply.

from non-combustion emissions). The allocation rule for splitting tier 1 between energy and non-energy purchases is the following: for each region x industry:

- if the industry is not an energy generation industry, all the impacts are considered non-energy (Share\_non\_energy = 1, Share\_energy = 0)
- if the industry is an energy generation industry, the impacts are split between energy and non-energy according to the share of GHG emissions due to combustion in the total GHG emissions of the industry (Share\_energy =  $\frac{\text{Combustion\_emissions}}{\text{Total\_emissions}}$ , Share\_non\_energy =  $\frac{\text{Non\_combustion\_emissions}}{\text{Total\_emissions}}$ ).

```

COMPUTE THE SHARE OF ENERGY TO NON ENERGY RELATED EMISSIONS FOR SCOPE 2
share_energy_non_energy <- D_CC %>%
  # Join with exiobase_industries to get the name (and not the ID) of industries and
  # allow string searches on industry names
  left_join(exiobase_industries, by = "ID_industry") %>%
  # identify electricity production industries
  mutate(Is_electricity=ifelse(str_detect(Exiobase_industry,"Production of electricity"),
    1,0),
    Is_steam = ifelse(str_detect(Exiobase_industry,"Steam and hot water supply"),
    1,0),
    # Identify energy related industries
    # Is_energy_industry equals to 1 or 0 as an industry cannot both be electricity
    # and steam related
    Is_energy_industry = Is_electricity + Is_steam,
    # the share of impacts allocated to energy is that of the share of combustion
    # emissions in total GHG emissions
    # for non-energy industries this share is 0
    Share_energy = Is_energy_industry *
    Combustion_emissions_per_mEUR/(Combustion_emissions_per_mEUR+Non_combustion_emissions_per_m
    EUR),
    # For energy industries, the share of impacts allocated to non-energy is that of
    # the share of non-combustion emissions in total GHG emissions
    # for non-energy industries this share is 1
    Share_non_energy = 1-Share_energy)

```

The Scope 2 impacts are then computed using the A matrix (to identify the direct purchases related to the turnover in each {region; industry}), D<sub>CC</sub> matrix (to get the GHG emissions caused by the production of the inputs purchased) and the share\_energy\_non\_energy matrix to split the emissions between Scope 2 and Tier 1:

$$\text{impact Scope } 2_j = i(M \circ D_{CC}) \cdot A \cdot x_j \times \text{share}_{\text{energy}}$$

In a future version of the tool, we might want to also attribute part of the Tier 1 LUEFN impacts to Scope 2, for instance to account for the land occupation of energy plants.



**Upstream Scope 3** impacts refer to all the remaining upstream impacts. They are computed based on the purchases related to the company's activity in the various {region; industry} pairs and the different suppliers' tiers (tier 1, tier 2, etc. suppliers) are distinguished thanks to the IO tables. Tier 1 impacts are computed based on the A matrix, while upstream value chain impacts are computed based on the Leontief inverse matrix (cf 2.1.A)

$$impact_j = i(M \circ D)(I - A)^{-1}x_j \sim impact_j = i(M \circ D)L.x_j.$$

```

COMPUTE VALUE CHAIN IMPACTS
# compute value chain impacts for each row of impact_data (a dataframe with pre-treated
activity production data and Scope 1 impact per pressure)
for(i in 1:nrow(impact_data)){
  row <- impact_data[i,] # select the row
  # get the corresponding column in the Leontief inverse matrix to compute the total input
purchases per mEUR
  vector <- L_inverse_matrix_integer[, (163 * (row$ID_region - 1) + row$ID_industry)] /
10^6
  # compute the purchases for the turnover
  purchases <- row$Amount * vector
  # compute the related impacts for each pressure. "*" replaces any pressure subscript
  impacts_* <- D_x_M_* * purchases
  # bind the results and handle column names
  # [...]
  # filter out the lines for which the total impact is null to limit the size of the
databases
  # [...]
  # store the impact of each line in a list
  impact_list[[i]] <- impacts
}
# Combine the elements of the list
impact <- do.call(rbind.data.frame, impact_list)

```

The Tier 1, Scope 2 and Scope 1 impacts are subtracted to the upstream value chain impacts in another part of the code to get the "rest of (upstream) Scope 3" impacts.

## 2.2 Using activity data to assess impacts

### A DIMENSIONING THE IMPACT

#### 2.2.A.1 Data pre-treatment

The computation of a default corporate footprint requires very limited data, namely the **turnover of the company broken down by region and industry of operation**. Ideally, those data should be provided in the **EXIOBASE nomenclature**, *i.e.* using EXIOBASE 3 region and industry terminologies (listed in (CDC

Biodiversité 2019a)), and **grouped by {region; industry} pairs**, *i.e.* splitting the turnover made in industry X between {region A; industry X} and {region B; industry X}. We learnt from the case studies that such detailed data is however seldom available. Most often, the data provided either do not fit the EXIOBASE nomenclature or document the region and industry mix separately instead of by {region; industry} pair. We thus apply data transformation rules so that the data format fits the required GBS input. Two guiding assumptions shape these rules. The first one is that getting industry figures right matters more than getting regional figures right because we assume that in most cases, difference between industries are more significant than regional differences (*e.g.* Cultivation of wheat is more different from Petroleum refinery than Cultivation of wheat in France is different from Cultivation of wheat in Germany). The second one is that the best hypothesis when data on turnover breakdown are lacking is to consider the average region or industry breakdown. In some situations, these assumptions may be clearly misleading, the assessors should then either take more appropriate assumptions, or avoid assessing the problematic entity altogether, if no satisfying assumptions can be made. Overall, assumptions should ensure assessments are fit for purpose, rigorous and consistent.

The cases most commonly encountered so far and the corresponding rules are presented below. They are applied through manual, and automated data treatment transcribed in the function `FI_pre_treatment`.

### 2.2.A.2 Dealing with data gaps

The procedure used to deal with data gaps, summarised by Figure 5, is explained hereafter.

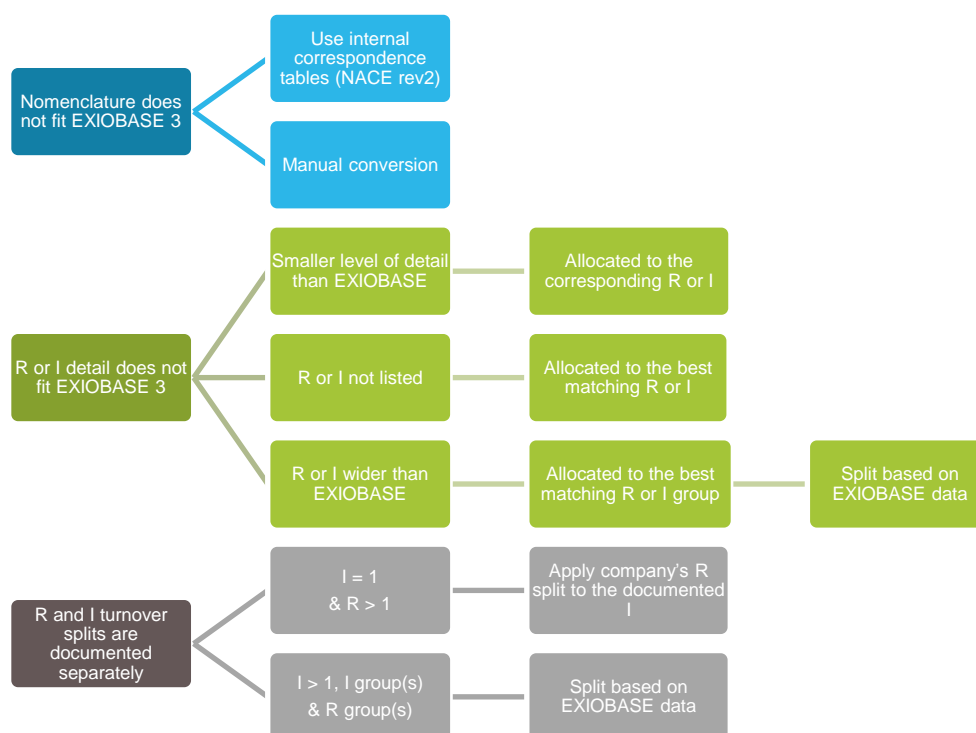


Figure 5: Dealing with data gaps. R = region, I = industry

635 ➤ **The data is in a nomenclature other than EXIOBASE 3.**

636 We established correspondence tables between the European NACE rev 2 nomenclature of industries and  
 637 EXIOBASE 3 (cf. (CDC Biodiversité 2019a)). Industry-related data can thus be provided in NACE rev 2  
 638 nomenclature, as well as in the French INSEE nomenclature. If the data is provided in another nomenclature,  
 639 we convert it manually to the most appropriate EXIOBASE category.

640 ➤ **The level of detail of the region or industry documented is different from that of EXIOBASE**  
 641 **categories**

642 As presented, EXIOBASE region categories are at the country or group of country level. Three sub-cases  
 643 can occur.

644 First, data can be provided at a lower geographical level than EXIOBASE (e.g. infra-national state or county).  
 645 They are then allocated to the corresponding country.

646 Second, data related to countries not individually listed in EXIOBASE are allocated to one of the  
 647 corresponding “Rest of” regions (for instance “Rest of Asia”).

648 Third, data related to wider geographical areas than the EXIOBASE 49 “regions” (e.g. a turnover reported  
 649 for the entire European Union and not separately for each Member State) are allocated to one of EXIOBASE’  
 650 11 region groups that fit the most commonly used regional entities (e.g. European Union, Asia, South  
 651 America...). If none of the 11 region groups fit, it is associated to the “World” region group, specific to GBS  
 652 analyses.

653 In this third case where turnover is specified at the region group level, data undergoes a second step during  
 654 which the associated turnover is split between the regions of the group based on the share of the production  
 655 of the group in each region. For instance, if the company operates in the industry “Cultivation of wheat” but  
 656 turnover data – say EUR 100 million – is documented for the European Union as a whole, the turnover will  
 657 be split between the countries of the European Union according to their share in the production of the  
 658 industry “Cultivation of wheat” in the total industry production in the European Union as reported in  
 659 EXIOBASE 3. Thus, we will assume that only 5% of the “Cultivation of wheat” turnover of the company is  
 660 made in Belgium if the Belgian production of wheat represents 5% of the European Union’s production of  
 661 wheat. In short, when data is documented at the group level, we assume that the company’s mix fits  
 662 EXIOBASE mix.

663 In all three cases, the rules applied to regions are similarly applied to industries.

664 ➤ **The region mix and industry mix of turnover are documented separately instead of by**  
 665 **{region; industry} pair.**

666 Three cases are distinguished.

First, if the company operates in only one industry and several regions, the turnover is split between the regions documented based on the share of turnover made in each region

Second, if the company operates in only one industry and one or more region groups, the turnover is first split by region group based on the share of turnover made in each region group to obtain {region group; industry} pairs. Then, it is split between the regions of each group based on the share of the region in the region group production for the industry to obtain {region; industry} pairs.

Third, if the company operates in several industries or industry groups and one or more regions or region groups, maintaining both the documented region and industry mixes while using average turnover breakdown (in line with the second guiding assumption listed above) would lead to inconsistencies. Considering the first guiding assumption that industry level data are more discriminating than region level data, we only use the company's industry level mix data and rely on EXIOBASE data to split the turnover between the regions listed.

➤ **The turnover split per region or region group is not available.**

We use EXIOBASE region mix for the industry of interest. For each industry, the turnover is split between the 49 regions according to the share of each region in the world production of the industry in EXIOBASE data and the rules described above are applied.

➤ **The turnover split per industry or industry group is not available.**

We manually classify the company into the most relevant EXIOBASE industry and consider that 100% of its turnover is made in this industry. If this assumption is not satisfactory, it might be preferable to keep the company out of the assessment altogether. The rules described above are applied.

## 2.2.A.3 Impact evaluation

### 2.2.A.3.1 Methodology

Impact evaluation is done through the functions `IO_evaluator_activity` (computes the impacts per pressure, scope and commodity, cf. above) and `IO_activity_assessment` (results formatting to display summary tables and graphs). Aside of impact tables, the default assessment also returns a "default inventory" (Scope 1 and whole value chain) as computed with the environmental extensions of EXIOBASE. The default inventory displays the quantity of raw materials (crops, non-primary crops, metals; in tons) and water (in Mm<sup>3</sup>) consumed due to the company's activity. When possible, these default inventories can be compared to companies' real inventories to investigate the potential differences between default and refined impacts, as well as to better understand default impacts. Default inventories are computed thanks to the function `how_many_commodities` that relies on D matrices and companies' turnover per {region; industry} to compute the corresponding amounts of raw materials.

```
COMPUTE DEFAULT INVENTORY FOR A SPECIFIED COMMODITY TYPE
how_many_commodities <- function(data, detail_level = NULL, commodity){
```

```

701
702   # The parameter "data" is either the pre-treated production data (for Scope 1 default
703   inventory) or impact data (output of the evaluator subfunction for value chain default
704   inventory). The column used is the one that indicates production (Amounts or Purchase, in
705   mEUR)
706   # If default inventory for the value chain, do required data manipulations
707   # [...]
708
709   # For the specified commodity c
710   if(commodity == c){
711     commodity_account <- data %>%
712       # Compute the quantity of the commodities of the type for each line based on the
713       # corresponding D matrix
714       left_join(D_LUEFN_*, by = c("ID_region", "ID_industry")) %>%
715       mutate(raw_materials_consumed_tons = Quantity * Amount)
716
717     # sum at the desired level of detail if required
718     if(is.null(detail_level)){
719       commodity_account <- commodity_account %>%
720         summarise(raw_materials_consumed_tons = sum(raw_materials_consumed_tons,
721                                                     na.rm = TRUE)) %>%
722         filter(raw_materials_consumed_tons != 0)
723     }
724     else {
725       commodity_account <- commodity_account %>%
726       group_by_at(detail_level) %>%
727       summarise(raw_materials_consumed_tons = sum(raw_materials_consumed_tons,
728                                                     na.rm = TRUE)) %>%
729       filter(raw_materials_consumed_tons != 0)
730     }
731   }
732
733   return(commodity_account)
734

```

Default inventories are computed only for the raw materials categories for which a CommoTool exists and are updated as new CommoTools are developed. Materials not yet taken into account in the assessment are also computed by the function `compute_materials_not_included` in order to give an idea of the size of the impact that is not accounted for. Materials not taken into account may be materials for which a CommoTool is under development (e.g. forestry) or materials for which no CommoTool is planned yet due to data limitation (e.g. fisheries).

#### 2.2.A.3.2 Example: biodiversity footprint of the production of EUR 1 million of French wheat

We illustrate the methodology using the example of the computation of part of the biodiversity footprint of the production of EUR 1 million worth of French wheat. The pressures assessed are listed in Figure 6. The whole upstream value chain is assessed, taking into account the impact of crops, metals and water consumption.

## Perimeter



Figure 6. Perimeter of the pressures assessed in this application (LUEFN: land use, encroachment, fragmentation and atmospheric nitrogen deposition)

The Wheat production industry in France purchases from many other industries and listing the impacts associated with all of them would make the example barely readable. In order to keep explanations simple and reader-friendly, the example thus highlights the impacts caused by two specific purchases, which represent only a small fraction of the total impacts.

The two steps of the default assessment computation described above are followed, the underlying data are provided in the package `GBStoolbox` in a `.rda` file called `example_IO_french_wheat_input_data.rda`. The application of the functions `IO_pretreatment_import()`, `IO_pretreatment` and `IO_evaluator()` gives the following results.

The **first step** links the production assessed to its direct biodiversity impacts. As explained, it is divided into two sub-steps. In the **first sub-step**, “direct environmental impacts” related to {France; Cultivation of wheat} are extracted from EXIOBASE environmental extensions. Here, the GHG emissions documented in the emission accounts is 842 943 kg CO<sub>2</sub>-eq. The raw material quantity documented in the material accounts is 4 080<sup>8</sup> tons of wheat.

In the **second sub-step**, biodiversity impacts are computed based on the inventories obtained through the first sub-step. For raw materials, the crop and metals CommoTools are used ((CDC Biodiversité 2020b), (CDC Biodiversité 2019e)). For pressures other than CC, the Scope 1 dynamic impact of the production of 4 080 t of wheat is assessed at 21 924 MSA.m<sup>2</sup>.

GHG emissions related impacts are computed using the “climate change” factor. The Scope 1 CC impact related to GHG emissions amounts to 3 684 MSA.m<sup>2</sup> for terrestrial biodiversity. The CC impact on aquatic biodiversity is 83 MSA.m<sup>2</sup>.

The **second step** repeats the analysis for suppliers. In this example, the perimeter is limited to direct (*i.e.* tier 1) suppliers. The amounts purchased for the production of EUR 1 million worth of French wheat are documented in the IO table. Among many other purchases, the purchase of EUR 7 500 from {Russia;

<sup>8</sup> In (CDC Biodiversité 2019b) this figure was 5 289. The difference is due to the fact that we now consider only the used extraction, as explained in 2.1.B.2.1.



Petroleum refinery} is required. In the rest of the text, we focus only on the impact related to {Russia; Petroleum refinery}. In the first sub-step, the GHG emissions are identified from the emission account (the raw material extraction induced could also be read from the material account but it is excluded from the perimeter of this example). The amount of refined petrol purchased causes the emissions of 5 794kg CO<sub>2</sub>-eq.

Concerning the Scope 2 impacts (impacts related to the generation of the electricity, steam, heat and cold purchased), IO tables provide all the non-fuel energy purchases. We focus on purchases from {France; Production of electricity by coal}. Only EUR 65 are purchased from {France; Production of electricity by coal}, inducing the emission of 732 kg CO<sub>2</sub>-eq.

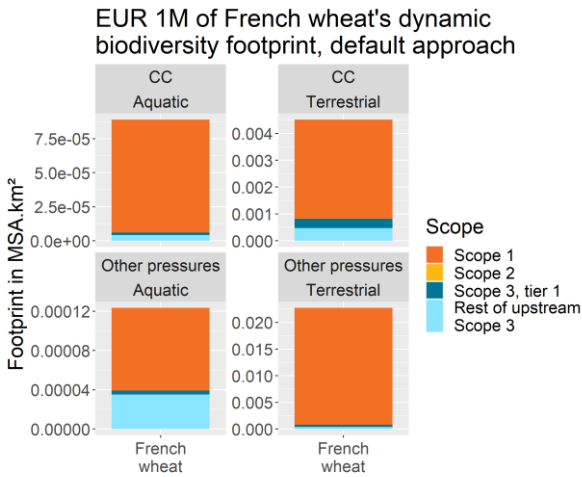
In the second sub-step, the biodiversity impacts for the two specific purchases we focus on are evaluated. For petrol purchases, they amount to 26 MSA.m<sup>2</sup> and for the Scope 2 impact of purchases from {France; Production of electricity by coal}, they amount to 3 MSA.m<sup>2</sup>. These impacts are very limited due to the perimeter of the study. In general though, Scope 2 and 3 represent a significant share of the total impacts of businesses, especially when impacts across the entire upstream value chain are considered.

The output of the assessment for French wheat is presented on Table 2<sup>9</sup> and Figure 7 to Figure 10. Static impacts amount to 7.8 MSA.km<sup>2</sup>, mainly occurring in Scope 1 due to terrestrial spatial pressures.

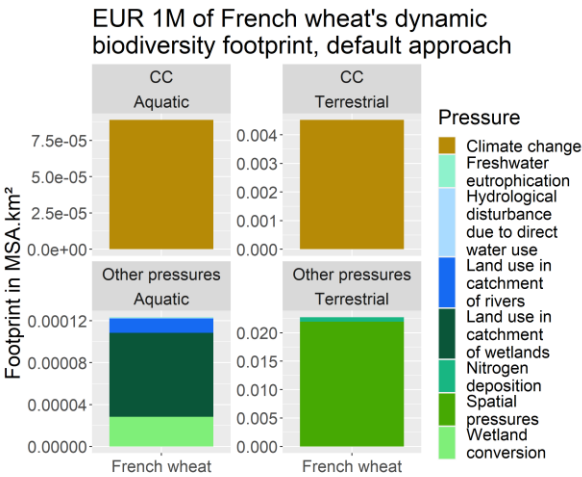
*Table 2: GBS output for the dynamic impact of the production of EUR 1 million worth of French wheat. CC: Climate change, Other pressures gather Land use, Encroachment, Fragmentation, Nitrogen deposition, HD: Hydrological Disturbance, Land use in catchment of rivers and wetlands, Wetland conversion, Freshwater eutrophication, TBA in 2020: To be added in 2020*

Perimeter	Scope 1		Scope 2		Scope 3, tier 1		Rest of Scope 3	
Pressure	CC	Other pressures	CC	Other pressures	CC	Other pressures	CC	Other pressures
Impact (MSA.m <sup>2</sup> )	3 683	21124	5.8	TBA in 2020	358	440	469	378

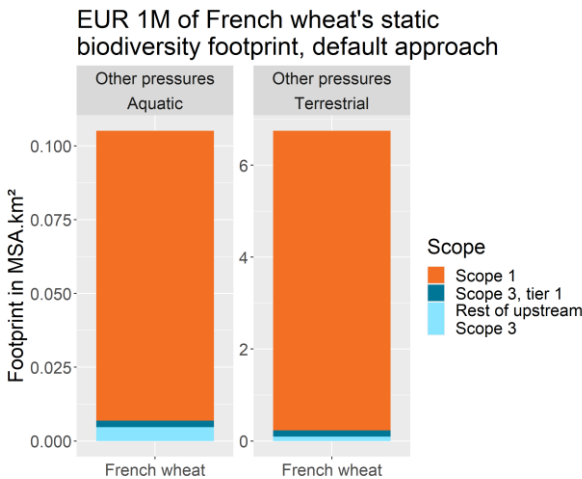
<sup>9</sup> Figures are different from the ones presented in (CDC Biodiversité 2019b) due to the addition of aquatic pressures. Also, the 3700 MSA.m<sup>2</sup> for Tier 1 – CC in this publication was a mistake.



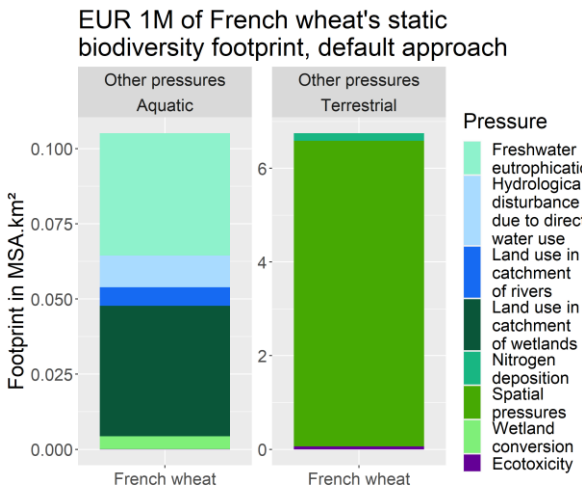
Source: GBS computation, Jul 2020



Source: GBS computation, Jul 2020



Source: GBS computation, Jul 2020



Source: GBS computation, Jul 2020

Figure 7: Default dynamic footprint of the production of EUR 1 million of wheat in France per scope

Figure 8: Default dynamic footprint of the production of EUR 1 million of wheat in France per pressure

Figure 9: Default static footprint of the production of EUR 1 million of wheat in France per scope

Figure 10: Default static footprint of the production of EUR 1 million of wheat in France per pressure

### 2.2.A.3.3 Tests

The results of the French Wheat example are used in the test `test-IO_french_wheat` of the package `GBStoolbox`. The expected results are saved in the `.rda` file `example_IO_french_wheat_expected_results.rda` in the folder `data` of the package. The test checks that computed and expected results are consistent. This is especially useful to check that new developments (e.g. new CommoTools and addition of the related impacts) do not alter computed impacts for older developments.

## B ATTRIBUTING

### 2.2.B.1 Financial asset focus: listed equities and corporate debt

This section deals with one of the three applications of the EEMRIO framework in the GBS: financial asset footprint assessment.

#### 2.2.B.1.1 Methodology

An equity portfolio can be seen as a bundle of business activities. Corporate loans can be seen in a similar light, with debt replacing equity. In both cases, the funding source can be considered to own part of the businesses it finances. Consequently, a part of the impacts generated by the businesses financed can be attributed to the funding source.

Assessing the **footprint of financial assets** involves two steps.

First, in the **dimensioning step**, the biodiversity impact of the business activities financed are assessed. If no specific data is available, it will involve the default corporate assessment methodology described above. At the current stage of development, this approach can be applied only to listed equity and to large corporation' debt. For private equity, data are usually too scarce to conduct default assessment with the methodology described above. For small and medium enterprise corporate loans, data will be similarly lacking. Specific methodologies will therefore need to be developed for those asset classes.

Then, in the **attribution step**, a fraction of the footprint of the companies financed is attributed to the funding source. To do so, attribution factors are computed for each company and are defined as the share of the company's enterprise value owned by the funding source, *i.e.*

$attribution\ factor_{company, portfolio} = \frac{financed\ value_{company, portfolio}}{enterprise\ value_{company}}$ . This is in line with the "share of the assets owned" approach to defining the perimeter under control (CDC Biodiversité 2020a).

The attribution factor is the same for listed equity and corporate loan. It is in line with the attribution factors used by the Platform for Carbon Accounting for Financials (PCAF 2017).

The total footprint is thus:

$$Footprint_{portfolio} = \sum_{company \in portfolio} footprint_{company} \times attribution\ factor_{company, portfolio}$$

Although the conceptual definition of the attribution factors is straightforward, practical issues occur when computing them in practice. Indeed, the value and the number of shares of the companies fluctuate over time, so that the attribution factors described above as  $\frac{\text{investment value}}{\text{enterprise value}}$  also fluctuate. For instance, let's consider the example we took in our Common ground working paper (CDC Biodiversité, ASN Bank, and ACTIAM 2018): a company Z with a debt of EUR 1000 and 10 shares with an initial value of EUR 100 per share and thus a market capitalization of EUR 1000. If the shares' valuation moves from EUR 100 per share to EUR 50 per share, the attribution factor changes:

Initial attribution factor for owner of 1 share:  $\frac{100}{10 \times 100 + 1000} = 5\%$ .

Attribution factor for owner of 1 share after the price change:

- If the value invested (EUR 100) is used:  $\frac{100}{10 \times 50 + 1000} = 6.6\%$

- If the current value of the investment is used (EUR 50):  $\frac{50}{10 \times 50 + 1000} = 3.3\%$

Also, the attribution factor changes if there is a share buy-back or a share emission, or if the ratio of the investment over the market capitalization evolves. Hence, assessing the attribution factors on a particular date, e.g. December 31<sup>st</sup>, may lead to biases. Computing attribution factors more frequently and averaging them over the period considered is a possible solution to this issue, though more data intensive. We recommend that the enterprise value be computed as the annual average of daily enterprise values at closure.

#### 2.2.B.1.2 Example: fictitious portfolio

The default assessment is illustrated on a fictitious portfolio gathering companies operating in food (Food 1 and Food 2), oil (Oil) and mining (Mining 1 and Mining 2) sectors. The portfolio is made up of several funds and portfolios owned by two asset managers (AM 1 and AM 2), its architecture is reproduced on Figure 11. The size of the portfolio is EUR 6.9 billion financed, while the companies gather a total turnover of EUR 75.8 billion (see

Table 3). The related attribution and dimensioning data are gathered in the file `example_I0_portfolio_input_data.rda`, the results are presented on Figure 12 to Figure 17. Figure 12 to Figure 15 show the aggregated results for the whole portfolio, while Figure 16 and Figure 17 present the result per company.

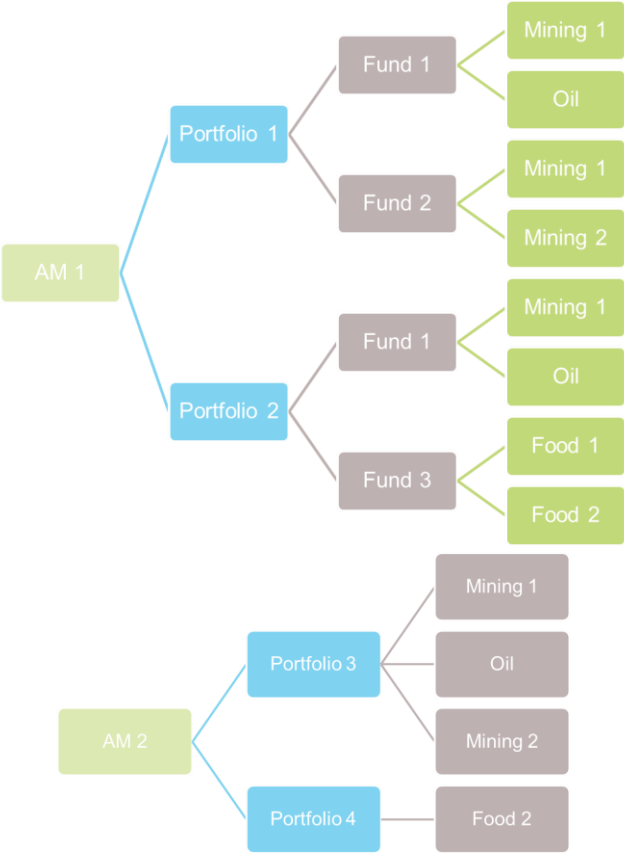


Figure 11: Structure of the fictitious example portfolio

859 Table 3: Composition of the example financial data

Portfolio	Fund	Company	Enterprise value (billion euros)	Total turnover (billion euros)	Share of total market capitalisation and debt in fund
Portfolio 1	Fund 1	MINING1	126.8	34.1	0.73%
		OIL	130.9	171.5	0.87%
	Fund 2	MINING1	126.8	34.1	0.37%
		MINING2	97.1	40	1.13%
Portfolio 2	Fund 1	MINING1	126.8	34.1	0.73%
		OIL	130.9	171.5	0.87%
	Fund 3	FOOD1	265.7	24.7	1.42%
		FOOD2	229.4	76.8	1.95%
Portfolio 3		MINING1	126.8	34.1	1.46%
		OIL	130.9	171.5	0.87%
		MINING2	97.1	40	3.37%
Portfolio 4		FOOD2	229.4	76.8	1.95%



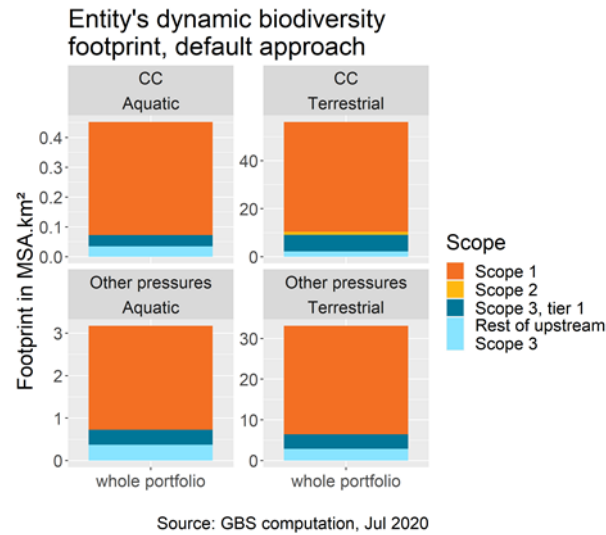


Figure 12: Example portfolio default dynamic footprint, per scope

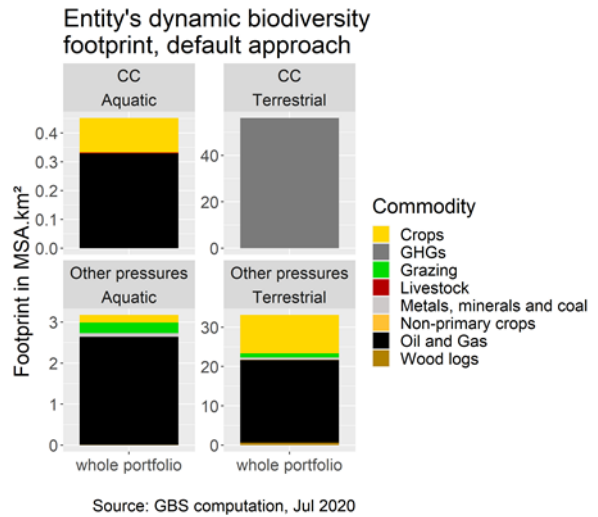


Figure 13: Example portfolio default dynamic footprint, per commodity type

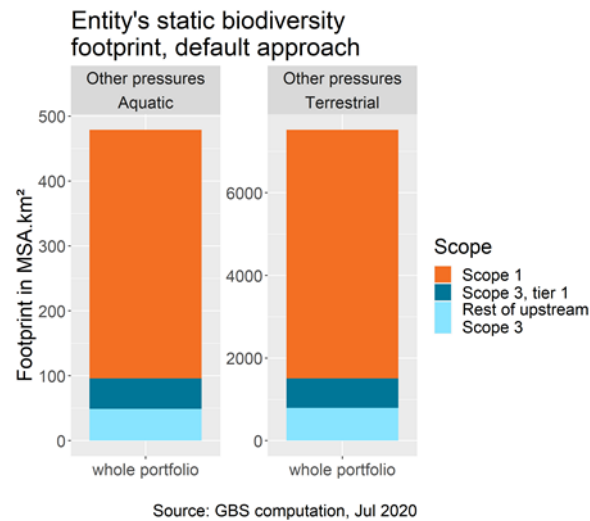


Figure 14: Example portfolio default static footprint, per scope

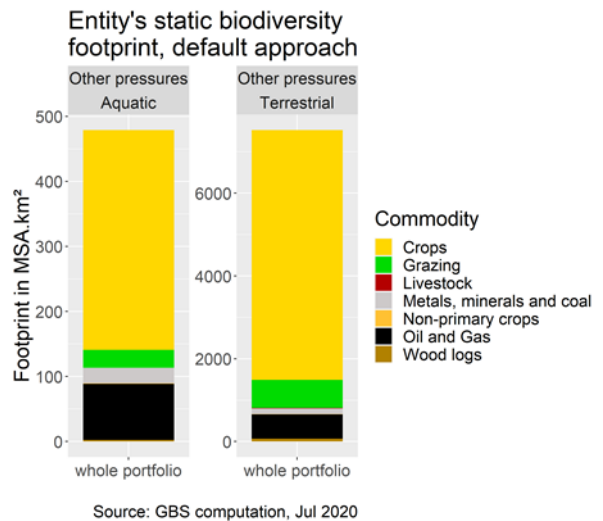


Figure 15: Example portfolio default static footprint, per commodity type

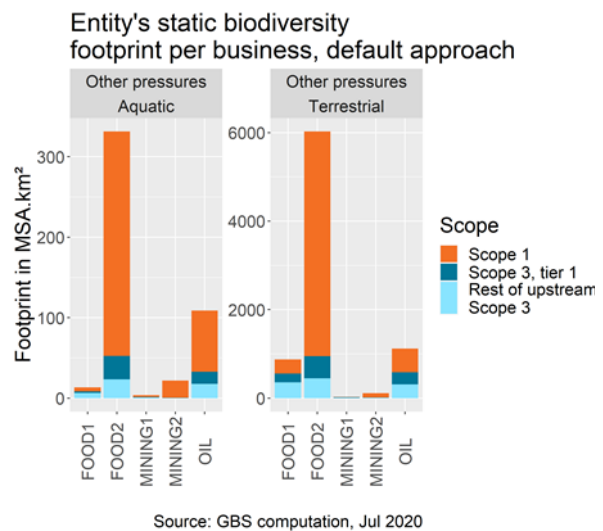
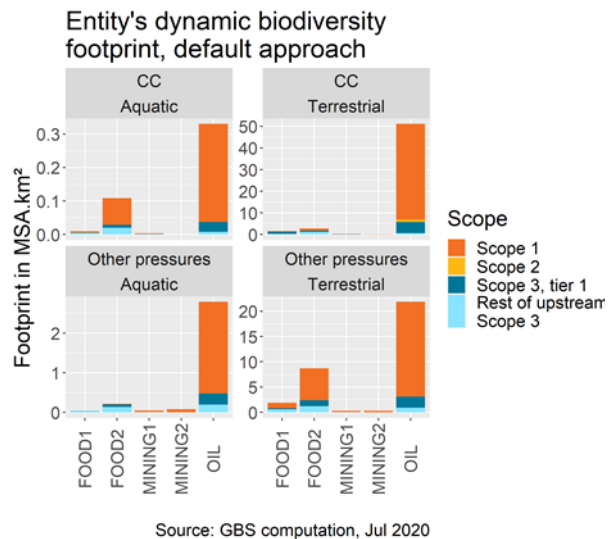


Figure 16: Example portfolio default dynamic footprint, per scope and company

Figure 17: Example portfolio default static footprint, per scope and company

860

861 The dynamic impacts are concentrated on terrestrial biodiversity, while static impacts are more evenly split  
 862 between terrestrial and aquatic biodiversity. The dynamic impacts amount to approximately 90 MSA.km<sup>2</sup>,  
 863 60% of which are due to climate change (CC) and 30% to the other pressures. Over 75% of the dynamic  
 864 impacts not related to CC occur in Scope 1, which is not surprising considering that companies in the  
 865 portfolio operate mostly in industries related to raw material production. Dynamic impacts are mainly due to  
 866 oil and gas, underlying the more important land use changes induced compared to metals and minerals.  
 867 Static impacts are predominantly due to crops and mining commodities, depicting the large surface areas  
 868 required by production. The static impact due to ecotoxic substances is reported separately and amounts  
 869 for the whole portfolio to 637 MSA.km<sup>2</sup> for terrestrial biodiversity, and 0.16 MSA.km<sup>2</sup> for aquatic biodiversity

870 Figure 16 and Figure 17 reveal that Oil is the company with the highest impact, due to the large GHG  
 871 emissions related to its activities. The second most impacting company is Food2, which has by far the  
 872 highest static impacts. The IO framework enables the computation of the Scope 1 and Scope 3 commodity  
 873 inventories related to Food2 activity to investigate further. Food2 operates in the industries “Cultivation of  
 874 crops, nec” in Indonesia and “Processing of food, nec” in Switzerland. The Scope 1 inventory thus relates  
 875 essentially to its cultivation activities in Indonesia while the Scope 3 inventory is more related to its food  
 876 processing activities in Switzerland. Scope 1 inventory contains a limited number of crop items displayed  
 877 on Table 4, as well as water consumption (10 Mm<sup>3</sup>) and various ecotoxic substances (91 188 kg). The  
 878 Scope 3 inventory is much larger, counting 200 items gathering crops, fodder crops, metals and minerals.  
 879 The impact of each item can be retrieved through their corresponding biodiversity impact factors, allowing  
 880 a precise analysis of the companies’ footprint.

These inventories correspond to economic activities as described in EXIOBASE v3, thus they may not match the company's real inventories. When companies can provide data on their real production/purchases (refined assessment), the related impact is computed and used in the results. For instance, if Food2 was able to provide its real Scope 1 inventory stating that it produces only 20 000t of Cocoa beans and 50 000t of Coffee, we would compute the impact of these productions and would use this impact instead of the impact of the inventory presented on Table 4. Details on refined assessments involving real purchases and inventory data are provided in Section 3.

Table 4: Food2 crops inventory, Scope 1

Commodity	Quantity (tons)	Commodity	Quantity (tons)
Natural rubber	164 413	Cinnamon	4 965
Cocoa beans	39 160	Pepper	4 789
Coffee, green	35 113	Cloves	3 975
Tobacco leaves	11 800	Nutmeg, mace, cardamoms	1 089
Tea	8 259	Spices, nec	275
Ginger	5 209	Vanilla	192

### 2.2.B.1.3 Tests

The results of the portfolio example are used in the test `test-IO_portfolio` of the package `GBStoolbox`. The expected results are saved in the `.rda` file `example_IO_portfolio_expected_results.rda` in the folder data of the package. The test checks the pre-treatment function and the structure of the overall pre-treated data. It also checks that computed and expected results are consistent. This is especially useful to check that new developments (e.g. new CommoTools and addition of the related impacts) do not alter computed impacts for older developments. The example also provides data to tests impacts aggregation rules.

## 3 Refined assessments

### 3.1 Using company data on purchases

In the IO framework, upstream value-chain impacts are computed based on the purchases documented by the IO model and computed using the Leontief inverse matrix. This procedure gives the monetary purchases

required to produce EUR 1 million of output of any {region; industry}. Of course, if the company assessed is able to provide monetary data on its own purchases, company data will be used instead of the default purchases as illustrated by Figure 18. This might require some data treatment to ensure that company purchases match the EXIOBASE nomenclature and it is best when monetary purchase data are provided by {region; industry} pair. Otherwise, pretreatment similar to the procedure explained in Section 2.2.A.2 is undertaken.

Theoretically, using companies' real purchases will yield more precise impacts. Yet, this is true only if the material flows related to the real purchases is closer to companies' real material flows than the flows computed based on IO purchases. This may not be the case. Let's take the example of Food2 involved in the fictitious portfolio in Section 2.2.B.1.2. Food2 operates in Switzerland in the industry "Processing of food, nec". Suppose that Food2 real purchases amount to EUR 50, while the purchases in the IO model are EUR 100. Then the material flows computed thanks to Food2 purchase data would be half those computed based on IO purchases, but both inventories would involve the same items (e.g. avocados and apples, even though the company possibly buys only apples). **Relying on real purchase data aligns monetary flows but does not allow screening in the materials flows embedded in them. The material flows computed based on real purchase data will match the company's real material flows only if 1) the company itself fits the average company in its industry in EXIOBASE and 2) the price structure matches that of 2011 (EXIOBASE data year).** Condition 1 implies that the company indeed buys all the items listed in the flows. Condition 2 ensures that the relative amount of each item is correct (see Section 4.2 for more details). These two conditions are likely not verified. Hence, relying on real inventory data is preferred.

## 3.2 Integrating refined assessments from inventories, pressures, direct measurements

As explained in (CDC Biodiversité 2020a), the GBS follows a hybrid approach to assess the footprint of economic activities. At any stage of the assessment, default data (based on IO model industry regional averages) can be replaced by real company data, either monetary (purchases, see Section 3.1) or not. Inventory data (tons of materials purchased, GHG emissions) can be used to replace default IO inventory data, while impacts can be computed directly based on pressure data (land occupation and land use type on operated areas). Of course, direct measurements can be integrated instead to replace computed impacts. Any refined value replaces the default value of the corresponding scope (static or dynamic) and pressure, while other default values are kept unchanged.

Figure 18 illustrates how the GBS proceeds through a stepwise approach. A less accurate default value is calculated first through the default assessment, then the parts of the footprints which can be recalculated through refined assessments are replaced. This work is handled by the function `IO_analyzer`.

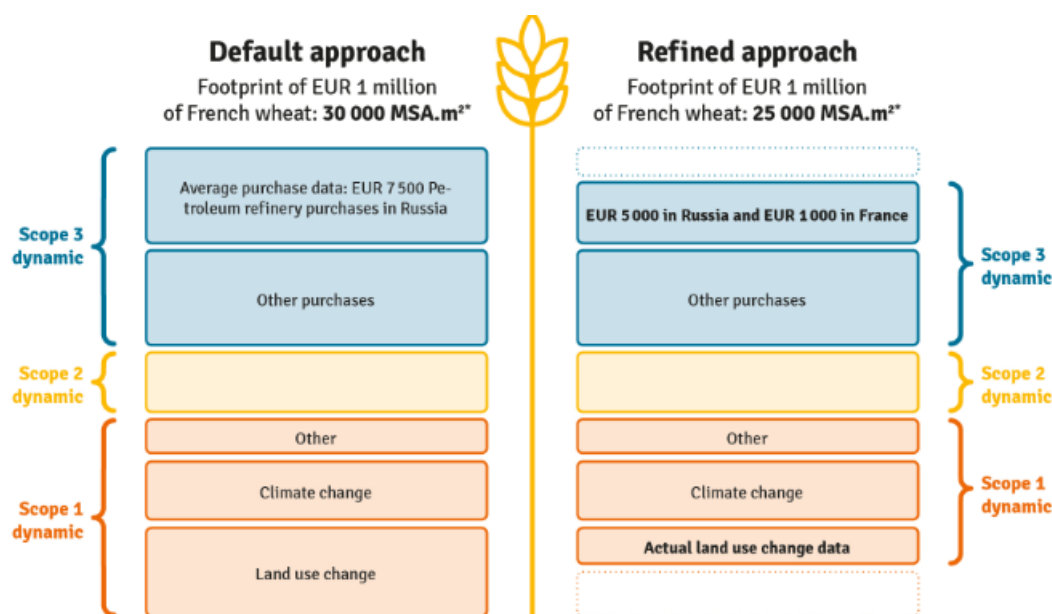


Figure 18: Illustration of the step-wise approach: less accurate footprints are replaced by more accurate ones when better data is available

## 4 Limits and perspectives

EXIOBASE EE MRIO database provides a consistent framework for tracking emissions, resource use and other environmental pressures along global supply chains. The combination of high sectoral detail and a wide range of environmental data is **fit for the purpose of the IO part of the GBS tool, i.e. describing corporate contribution to drivers of biodiversity loss within a spatialized framework considering both their direct contribution and that of their supply chains**. Several limitations of the EXIOBASE model and data, as well as limitations related to their use in the GBS tool should however be acknowledged.

### 4.1 Limitations of EXIOBASE 3

Building an EE MRIO database is a data-intensive task. Indeed, it requires to combine a heavy load of data of various types, from monetary supply-use tables stemming from national accounts to energy balances, trade data, modelled emissions factors and material accounts. Such a work of building consistency between numerous elements cannot be completed without assumptions regarding data correspondence, formats,

aggregation and allocation. The choices of EXIOBASE modellers is however precisely documented in several Supporting information materials.

Because building MRIO datasets requires important data work, a trade-off often exists between regional and sectoral detail. Namely, MRIO models either provide a wide range of industries and a limited number of regions, or the opposite: fewer industries but numerous regions. EXIOBASE 3 falls into the first category. Hence, the regional detail is sometimes scarce, especially for the five “Rest of World” regions. This is all the more important for our study given that such regions are often critical when it comes to biodiversity impact assessment. Sectoral precision seemed however even more important, hence our choice to use EXIOBASE 3 rather than a more regionally detailed MRIO database.

Although very detailed, EXIOBASE environmental extensions also suffer from limitations, not in terms of items listed but rather in terms of sectoral allocation. Indeed, some extensions like the water account only concern a limited number of industries, in this example Agriculture, Livestock breeding, Manufacturing and Electricity related industries. This causes default assessments by the GBS to endogenously ignore part of companies impacts as their contribution to some pressures will not be accounted for. This part can be material, for instance for the industry “Collection, purification and distribution of water”. This fact reinforces the need to complement IO default data by company data whenever possible.

## 4.2 Limitations regarding the use of IO modelling in the GBS and perspectives for future developments

Limitations related to the use of sectoral averages and incomplete inventories and the way these issues are dealt with within the GBS have already been presented.

The first limitation of the default assessment procedure in GBS relies on a theoretical question. As explained, default assessments consist basically in

1. Using turnover data to associate the company’s activity to the right industries and regions in the EXIOBASE model
2. Use the data in the environmental extensions (GHG emissions, raw material consumptions...) and our in-house developed tools to deduce the biodiversity impact

Step 1 is very important since it determines the accuracy of the computed value, so we want to be as close to the company’s real activity as possible. Hence, we want to get the industry(ies) and region(s) right. Companies’ annual reports often document the turnover (and not the production) split by countries/regions and by the company’s main sector of operation. The latter split can be used to split the company’s turnover between the right EXIOBASE industries, as explained in Section 2.2A. In our understanding EXIOBASE IO tables are based on **sales** (i.e. turnover: where were products and services sold? where are the customers?), while the environmental extensions are based on **production** (where were the goods and



services produced, where are the factories). As the turnover regional split – *i.e.* where the sales occur – may not be a good proxy for where production took place, we wonder whether using the turnover split is the best way to proceed. It came out from discussions with other developers that doing so is common, production data being much less documented. **The choice made for now remains however open and we might consider changing it if future discussions advise so.**

Another limitation is certainly related to the date of the data. Indeed, the most recent dataset provided by EXIOBASE v.3 is for year 2011. The time gap between 2011 and 2020 (year of the first GBS full corporate and portfolios assessments) is non-negligible. It is a concern in terms of results accuracy since **the structure of the economy, global prices and relative prices changed between the two periods.** The preferred way to solve this issue is the update of the IO data used in GBS when EXIOBASE dataset for more recent years become available. EXIOBASE developers announced that data for the year 2016 will be released soon. The time gap arguments further in favour of the use of company specific data whenever possible. In this regard, the flexibility of GBS in terms of input data used presented in (CDC Biodiversité 2020a) and in Section 3.2 of this report is thus critical and an undeniable strength of the tool. **Also, a possible way to reduce the bias related to time gap that we would like to discuss with experts and explore in the future would be to translate companies' activity data expressed in today's currencies into 2005 euros, the unit used in EXIOBASE 3. Indeed, this translation might limit the bias related to changes in prices structure between to two periods.**

Linked to the previous point, the fine resolution of some extensions compared to EXIOBASE industry resolution induces biases. Indeed, the material extensions related to crops are much finer than the Cultivation industries; only wheat and rice production sectors are singled out, causing for instance all the fruits, vegetables and nut to be produced by the same industry "Cultivation of vegetables, fruits and nuts". In the GBS, monetary flows are always translated into raw material flows. Thus, euros of production or purchases are transformed into tonnes of apples and avocados for instance, and this number of tonnes is different for the two products. Combined with the material account, the IO table thus provides a "price vector" for all products extracted by the industry. This price vector is fixed and corresponds to the year of the data (2011). Hence, **changes in the material flows or relative prices of products extracted by an industry cause the inventories computed through the IO model, and thus corresponding impacts, to be flawed.** For instance, if the real price of one of the items in the inventory changes, the impact computed through the IO model would remain the same since the change is not reproduced in EXIOBASE data. Therefore, the computed impact will be overestimated (the price increase causes less tonnes to be bought, so that the true impact is smaller after the increase). Two options exist to overcome this limitation: replacing default inventories computed through the IO module by real inventory data provided by the assessed companies (refined assessment, see Section 3) and working with more recent IO data. EXIOBASE will soon release its 2016 update, which will be integrated into the GBS. In our opinion, EXIOBASE sectors provide acceptable details of the products value chain, *i.e.* distinguishing raw material extraction from processing, manufacturing, distribution, waste treatment/recycling etc. Yet, it is true that:

1. Production of/purchases to an industry will always be considered to concern all the materials extracted by the sector (no possibility to buy only apples and no avocados);
2. Monetary flows between apple producers and avocados producers, for instance, cannot be isolated.

Hence, we insist that IO data are simple averages, so that **the use of IO modelling only provides a sectoral benchmark and default value of a company's footprint.** If a company can provide real data related to its



production and purchases (for instance showing that it purchases only apples), then this data and the related impact are used to replace the default value in the results (the default impact of the production of avocados is set to 0). Also, the **IO framework is more suited to the assessment of large companies, which activities better match the delineation of EXIOBASE industries**. As such, small and medium enterprises (SMEs) are currently not covered by the GBS. Yet, the approach used for large companies could be applied to SMEs. Two main questions would then need to be solved: 1/ how to attribute impacts of SMEs to capital owners (the enterprise value and the share of enterprise value owned may be unknown)? And 2/ is the impact sufficiently large for the GBS estimation to be reasonably accurate? The conditions required for that are described in (CDC Biodiversité 2020e).

## References

- BGS, - British Geological Survey. 2014. 'World Mineral Statistics'.
- CDC Biodiversité. 2017. 'Global Biodiversity Score: Measuring a Company's Biodiversity Footprint'. 11. Biodiv'2050 Outlook.
- . 2019a. 'GBS Review: Appendix'. Phase 1.
- . 2019b. 'Global Biodiversity Score: A Tool to Establish and Measure Corporate and Financial Commitments for Biodiversity'. 14. Biodiv'2050 Outlook. CDC Biodiversité.
- . 2020a. 'GBS Review: Core Concepts'.
- . 2020b. 'GBS Review: Crops CommoTool'.
- . 2020c. 'GBS Review: Freshwater Pressures on Biodiversity'.
- . 2020d. 'GBS Review: Mining CommoTool'.
- . 2020e. 'GBS Review: Quality Assurance'.
- . 2020f. 'GBS Review: Terrestrial Pressures on Biodiversity'.
- CDC Biodiversité, ASN Bank, and ACTIAM. 2018. 'Common Ground in Biodiversity Footprint Methodologies for the Financial Sector'. Paris: ACTIAM, ASN Bank, CDC Biodiversité. Supported by Finance in Motion. <https://www.asnbank.nl/web/file?uuid=b71cf717-b0a6-47b0-8b96-47b6aefd2a07&owner=6916ad14-918d-4ea8-80ac-f71f0ff1928e&contentid=2412>.
- IEA, International Energy Agency. 2014a. 'Energy Statistics of Non-OECD Countries. 1960-2011. Online Database'.
- . 2014b. 'Energy Statistics of OECD Countries. 1960-2011. Online Database'.
- PCAF. 2017. 'Paving the Way towards a Harmonised Carbon Accounting Approach for the Financial Sector'. The Netherlands: Platform Carbon Accounting Financials.
- Reichl, C., M. Schatz, and G. Zsak. 2014. 'World Mining Data'. *Minerals Production Inter-National Organizing Committee for the World Mining Congresses* 32 (1): 1–261.
- Stadler, Konstantin, Richard Wood, Tatyana Bulavskaya, Carl-Johan Södersten, Moana Simas, Sarah Schmidt, Arkaitz Usubiaga, José Acosta-Fernández, Jeroen Kuenen, and Martin Bruckner. 2018. 'EXIOBASE 3: Developing a Time Series of Detailed Environmentally Extended Multi-Regional Input-Output Tables'. *Journal of Industrial Ecology* 22 (3): 502–515.

- 1065 Stocker, Thomas. 2014. *Climate Change 2013: The Physical Science Basis: Working Group I Contribution*  
 1066 *to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge  
 1067 University Press.
- 1068 USGS, US Geological Survey. 2014. 'International Minerals Statistics and Information'.
- 1069 Wilting, Harry C., and Mark M. P. van Oorschot. 2017a. 'Quantifying Biodiversity Footprints of Dutch  
 1070 Economic Sectors: A Global Supply-Chain Analysis'. *Journal of Cleaner Production* 156 (July):  
 1071 194–202. <https://doi.org/10.1016/j.jclepro.2017.04.066>.
- 1072 ———. 2017b. 'Quantifying Biodiversity Footprints of Dutch Economic Sectors: A Global Supply-Chain  
 1073 Analysis'. *Journal of Cleaner Production* 156 (July): 194–202.  
 1074 <https://doi.org/10.1016/j.jclepro.2017.04.066>.
- 1075 Wilting, Harry C., Aafke M. Schipper, Michel Bakkenes, Johan R. Meijer, and Mark A. J. Huijbregts. 2017.  
 1076 'Quantifying Biodiversity Losses Due to Human Consumption: A Global-Scale Footprint Analysis'.  
 1077 *Environmental Science & Technology* 51 (6): 3298–3306.  
 1078 <https://doi.org/10.1021/acs.est.6b05296>.  
 1079

**CDC** BIODIVERSITÉ



102, rue Réaumur  
75002 PARIS  
T. +33 (0)1 80 40 15 00

[contact@cdc-biodiversite.fr](mailto:contact@cdc-biodiversite.fr)

[www.cdc-biodiversite.fr](http://www.cdc-biodiversite.fr)

SAS au capital de 17 475 000 euros  
RCS Paris 501 639 587  
Siret 501 639 587 00028 - APE 6420Z  
N° TVA Intracom. FR51501639587

**CDC Biodiversité**  
102, Rue de Réaumur  
75002, Paris – France  
T: (0)1 80 40 15 00

**Agence Sud-est**  
Bâtiment U1B  
1 B Allée du Square  
1140 Rue André Ampère  
13290 AIX EN PROVENCE  
T: (0)4 28 38 06 83

**Agence Sud-Ouest**  
333 boulevard du Président WILSON  
33000 BORDEAUX  
T:(0)5 32 09 07 91  
(plan d'accès)