

Master Thesis – End of course Internship Report

Multi Regional Input Output modelling as a tool for sustainable decision-making

February – August 2019

Timothé Beaufiles

beaufiles.timothe@gmail.com
Promotion 2019

Cursus Ingénieur Centralien, CentraleSupélec
Option Energie, Parcours Réseaux d'énergie (OEN)
Filière Métiers de la Recherche (FMR)

Master Economie de l'Energie, de l'Environnement et des transports
Parcours Modélisation Prospective : économie, énergie, environnement

Pedagogic Supervision

OEN : Hervé Duval
FMR : Pierre Lecoy
MEEET : Pascal Da Costa

Executive Supervision

PwC France: Etienne Berthet
PIK : Anders Levermann

université
PARIS-SACLAY


CentraleSupélec


pwc


P I K

ACKNOWLEDGMENTS

This report is the outcome of an intense internship, which gave me the chance to work in a stimulating environment with very inspiring people. First of all, heartfelt thanks Etienne Berthet for everything he did to settle this internship up and for his guidance all along this journey. I acknowledge that not a piece of this work would have been done without his enthusiasm and his determination, and I will never thank him enough for boarding me on that thrilling adventure. Thank you very much to Emilie Bobin and Sylvain Lambert for the risks they took and the confidence they put in such project. They offered me to work in exceptional conditions of autonomy, and I'm sure that their daring will pay back. Thank you to Olivier Muller for challenging me all along the project, and Jean-Baptiste Petit for the precious feedbacks on the tool. Thank you to Antoine Fesson for the brainstormings on the conception and the visualization of the tool, and to Caroline Devaux for her help. Thank you to Ludovic, Maxime, Rothmane and all my colleagues at PwC for their company and the fruitful discussions we had during the break times. I sincerely hope the outcome of my work there will make a difference on your way toward a sustainable future.

I am very proud that Anders Levermann agreed to host me at the Potsdam Institute for Climate Impact Research for a portion of my project. I warmly thank him for offering me a chance, as I thank once again Etienne for making this stay possible. I had a fantastic time there, having the opportunity to do stimulating science. I specially thank Leonie Wenz for mentoring me in the appealing project of forecasting the evolution of the world economy, along with Anders Levermann. Thank you to Anja Bruhn for making the administrative formalities so easy, and to Johannes Gütschow for his availability for the PRIMAP project. Of course, I would like to thank sincerely all my colleagues there: Kilian, Sven, Anne, Timon, Falk, Yao, Nicole, Nadine, Jan, for their kind welcoming, the fruitful discussions we had and the coffees after lunch. I had a great time there, and I can't wait to be back for the next steps! Also, I specially thank Sofia, Benedikt, Tim and Janusch for hosting me during all my time in Germany.

Thank you to Pierre Lecoy, Hervé Duval and Pascla Da Costa for following up this project and for their advices. Thank you to Didier Duval for his availability and his efficiency for the administrative formalities.

RESUME

La mondialisation et la spécialisation des acteurs économiques augmentent la longueur et la complexité des chaînes d'approvisionnement industrielles. La multiplication des interdépendances commerciales à l'échelle mondiale crée des défis majeurs pour les responsables politiques et économiques : elle augmente notamment la vulnérabilité aux réactions en chaîne dans les réseaux d'approvisionnement, et complique l'allocation des responsabilités des externalités socio-environnementales. Les modèles Entrée-Sortie Multirégionaux offrent une vision globale et systémique des interdépendances entre secteurs économiques au niveau mondial. Ce mémoire illustre en quoi ce cadre théorique offre une approche pertinente pour répondre aux défis de soutenabilité globale d'aujourd'hui et de demain. Grâce à leurs propriétés mathématiques, ils permettent ainsi d'appréhender les interactions sur lesquels reposent les chaînes d'approvisionnement. Le développement d'un outil basé sur ces principes pour PwC France souligne l'utilité de cette approche pour répondre à des défis opérationnels actuels. Enfin, le développement d'une nouvelle méthode analytique souligne le potentiel de ces modèles pour construire des scénarios prospectifs d'évolution de l'économie mondiale. Il apparaît ainsi que les modèles Entrées-Sorties présentent d'ores et déjà des applications commerciales pertinentes et laissent apparaître une marge de progrès importante pour appuyer des décisions stratégiques durables.

ABSTRACT

The globalization and the specialization of economic actors increase the length and complexity of industrial supply chains. The multiplication of interdependencies throughout the world economy triggers substantial challenges for decision makers: they increase the vulnerability to shockwaves throughout supply networks and blur the allocation of responsibilities for social and environmental externalities. Multi Regional Input Output modelling provides a complete and systemic overview of the interlinkages throughout the world economy. This thesis highlights how this framework offers a privileged approach to address sustainability related challenges. By their mathematical properties, Input Output tables permit allocating the externalities throughout supply chains. The development of an operational tool for PwC France underlines that this approach already fills some operational needs. Moreover, the design of a new analytical method underlines the potential for Input Output modelling to support the development of prospective scenarios for the world economy. Therefore, Multi Regional Input Output modelling already offers relevant applications to business concerns and show a substantial the room of improvement for supporting sustainable decision making.

TABLE OF CONTENTS

Acknowledgments.....	2
Résumé	4
Abstract	4
Table of Contents	6
Foreword	9
Introduction	10
Chapter I: Environmentally Extended Input Output Modelling, Theory and Practice	13
I.A. An introduction to Input Output Analysis.....	14
1) <i>The Input Output framework.....</i>	14
2) <i>Notations</i>	15
3) <i>Iterative approach</i>	16
4) <i>Leontief inverse.....</i>	17
5) <i>Application to primary input analysis.....</i>	18
6) <i>Structural decomposition</i>	19
7) <i>Limits and drawbacks.....</i>	20
8) <i>Synthesis.....</i>	22
I.B. Environmental Extension	23
1) <i>Presentation</i>	23
2) <i>Integrating environmental data to economic tables: merging PRIMAP and EORA</i>	25
3) <i>Study case: Consumption Based Accounting of global Greenhouse Gases Emissions.....</i>	33
Chapter II: A tool to assess the social and environmental footprint of supply chains.....	38
II.A. Overview	39
1) <i>Indicators</i>	39
2) <i>Treatment modes</i>	39
3) <i>Input data</i>	40
4) <i>Output</i>	40
5) <i>Documentation.....</i>	42
6) <i>Study case.....</i>	42
II.B. Data, reliability and comparability	46
1) <i>Economic data</i>	46

2) <i>Environmental data</i>	47
3) <i>Reliability</i>	48
4) <i>Compatibility with Life Cycle Analysis</i>	49
II.C. Technical Implementation	53
1) <i>Main codes</i>	53
2) <i>Data repositories</i>	55
3) <i>Interface</i>	56
4) <i>Delivery</i>	57
5) <i>Update routine</i>	57
II.D. Synthesis and perspectives	59
Chapter III: Forecasting plausible evolutions of the world economy	61
III.A. Contextualization	62
III.2. Method: an algorithm to project future trade relations	63
1) <i>Step 1: From macroeconomic projections to aggregated constraints</i>	63
2) <i>Step 2: Updating the matrix</i>	65
3) <i>Technical Implementation</i>	65
III.3. Validation of the performance of the algorithm	67
1) <i>First step validation</i>	67
2) <i>Second step validation</i>	68
3) <i>Holistic validation</i>	70
III.4. Discussion & further work	72
Conclusion	74
Abbreviations	76
Glossary	77
Bibliography	79
Annex I: Matching EORA economic sectors with 2006 IPCC's emissions categories	84
Annex II: Quality assessment of the EORA's environmental indicators	86

FOREWORD

This master thesis gathers the outcome of a 6-months internship running from February to August 2019. It took place into the Sustainability Performance Strategy's (SPS) division at PwC France, in Neuilly-Sur-Seine (France), under the direct supervision of M. Etienne Berthet¹ and included a 3-months stay at the Potsdam Institute for Climate Impact Research (PIK), at Potsdam (Germany) under the supervision of Pr. Anders Levermann². This report and the internship credited for the obtention of two master's degrees: for the French title of engineer at *CentraleSupélec*, in the field of energy grids with a special focus on research work; and for the master *Environnement, Energie et Transport*, in the *Modélisation prospective: Economie de l'énergie et de l'environnement* path, delivered by *Université Paris-Saclay*. M. Olivier Duval³ and M. Pierre Lecoy⁴ were in charge on the pedagogic follow-up of the thesis for the account of *CentraleSupélec*. M. Pascal Da Costa⁵ was in charge for *Université Paris-Saclay*.

The final version of this thesis was printed the 21st October 2019. It was orally supported on the 4th of November 2019 at CentraleSupélec⁶, in front of the following juries: M. Etienne Berthet on behalf of PwC France, M. Pascal Da Costa on behalf of *Université Paris-Saclay*, M. Hervé Duval and M. Pierre Lecoy on behalf of *CentraleSupélec*.

¹ etienne.b.berthet@pwc.com

² Anders.Levermann@pik-potsdam.de

³ olivier.duval@centralesupelec.com

⁴ pierre.lecoy@centralesupelec.com

⁵ Pascal.da-costa@centralesupelec.com

⁶ 8-10 rue Joliot-Curie – 91 190 Gif-sur-Yvette - France

INTRODUCTION

From the end of July 2011, the conjunction of the monsoon and a tropical storm triggered five months of severe flooding in Thailand, causing 815 casualties and the shutdown of many industrial plants in the country. This local event triggered chain reactions in the world economy, causing disturbances on many industrial sectors, including a surge in Hard Disk Driver prices all over the world, that lasted for several month (Haraguchi and Lall 2015). This example is not an exception: extreme weather events regularly create tensions on various markets all around the world (Levermann 2014), and illustrate the vulnerability of globalized supply chains to local shocks.

Simultaneously, the internationalization mobilization to tackle climate change requires a drastic diminution of carbon emissions at the world level (Le Quéré et al. 2018). The allocation of the responsibility for Greenhouse Gases (GHG) emissions is a burning question, on which depends the response to this global issue. Who really causes the emissions? Big companies which extract fossil fuels, or their consumers who burn them? The allocation of the responsibilities for Greenhouse Gases (GHG) emissions rises a crucial scientific challenge. How could we assign such externalities, when so many intermediate actors intervene between primary resource extractors and final demand consumers?

Both issues become increasingly critical for decision making. Unexpected disruption in supply chains may unpredictively endanger industrial business, while public opinion and regulatory institutions push economic actors to account for their externalities. In both cases, the key of the problem lies in the understanding of the interrelations throughout the world economy. Such knowledge is obviously hard to get. It is now normal business to have suppliers from other continents, which themselves source their inputs from other countries. The exponential complexity of those trade relations and their potential confidentiality impedes a local-scale approach. Manual investigation could hardly map two and three suppliers, and falls short for systemic propagation effects.

Multi Regional Input Output (MRIO) is a pivotal approach to this class of problem. Providing a system-wide overview of economic networks, such modelling offers otherwise unreachable insights on the functioning of economic networks. During 6 months of internship at PwC France's Sustainability Performance Strategy team and at the Potsdam Impact for Climate Impact Research (PIK), I had the chance to explore the potential of MRIO as a tool for sustainable decision making. Being at the intersection of commercial applications and theoretical research, I focused on applying state-of-art analysis techniques to meet practical challenges encountered by professional businesses. Simultaneously, my experience at PIK offered me the opportunity to deepen my understanding of MRIO models and to investigate their potential for prospective analysis. This report provides a detailed summary of the insights of this internship, by examining some high potential applications offered by MRIO.

The first chapter serves as a global introduction to the Input Output (IO) framework. I particularly focus on Environmentally Extended Input Output (EEIO), highlighting the specific forces, challenges and drawbacks embodied by this approach. While this chapter provides large parts of theoretical demonstration, it contains two pieces of original researches, with the integration of a GHG database into an IO table, and the analysis of the trade of GHG emissions between the major economies of the world.

The second chapter provides the detailed report of the development of an operational tool for PwC France, designed to assess the social and environmental footprint of economic activities. This application constitutes the main output delivered to PwC. It aims at filling the growing need for companies to understand their impact on the society, to improve their sustainability performance.

While the implementation of this tool relies on pre-existing science, I designed it to be as consistent and flexible as possible. It therefore includes an in-depth discussion on the reliability of such approach, and its compatibility with resembling techniques.

The third and last chapter constitutes an investigation toward extended use of MRIO tables for prospective studies. This fully original work includes the development of an analytical method to forecast MRIO tables in a frame common to the academic community of prospective modelling. Apart from the detailed presentation of this method, this chapter includes a detailed assessment of its performance, and draws perspectives toward further improvements and applications.

CHAPTER I: ENVIRONMENTALLY EXTENDED INPUT OUTPUT MODELLING, THEORY AND PRACTICE

Input Output Analysis (IO) dates from the Nobel-Prize laureate Wassily Leontief's work in the 1930s (Leontief 1936). First limited by computational limits and data availability, this approach attracted a growing attention during the last decades. The release of worldly consistent Multi Regional Input Output (MRIO) tables triggered the development of numerous applied methods in economics and in industrial ecology: value-added analysis (Aslam, Novta, and Rodrigues-Bastos 2018), footprint calculation for carbon emissions (T. Wiedmann 2009) and other socio-environmental externalities (Alsamawi et al. 2014; T. O. Wiedmann et al. 2015). This first chapter provides a short introduction to Environmentally Extended Input Output Analysis (EEIO). The first part introduces the basics of Input Output Analysis, and some practical aspects of its concretization. The second part focuses on its environmental extension.

I.A. An introduction to Input Output Analysis

Following paragraphs offer a basic presentation of Input Output Analysis, along with a critical study of some practical aspects. Reader could refer to Miller and Blair for a more detailed insight on the topic (Miller and Blair 2009).

1) The Input Output framework

In their simplest form, Input Output tables are two-dimension matrices of which each row and each column represent a different industrial sector. Rows contain the output of each sector and its destination, may it be to other sectors of the economy (intermediate goods) or to consumers (final goods), most often in monetary values. Conversely, columns represent inputs flowing into one sector. Once again, we distinguish intermediate inputs, from the other sectors of the world economy and primary inputs or value-added¹. In respect to basic accounting principles, there is monetary equivalence between the total output flow from one sector (gross output), and the respective total input flow (gross input). Normalizing each column by the total output of the respective sector gives the *Technical matrix*, where each coefficient is smaller than one and higher than zero. They represent the relative importance of one input (in row) in the production of another product (in column). The distribution of coefficient in a row therefore offers the 'recipe' to produce a unit of a given output.

	Farms	Bakeries	Final Demand
Farms	20	120	20
Bakeries	32	20	148

Capital	68	30
Labour	40	30

Illustration 1A - Input Output Table of an economy with two productive sectors. Farms produce wheat flour, and bakeries produce bread. Farms produce a total of 160 units of wheat flour. Bakers buy 120 units of it, and 20 go directly to final consumers. Farmers keep the last 20 units to plant next year's wheat. Bakers produce 200 units of bread. They sell 148 of it to final consumers, while they consume 20 for themselves, and sell 32 to the farmers.

	Farms	Bakeries	Final Demand
Farms	0.125	0.6	10
Bakeries	0.2	0.1	130

Capital	0.425	0.15
Labour	0.25	0.15

Illustration 1B – Technical matrix associated with illustration 1A. In order to produce one unit of bread, bakeries require 0.6 units of wheat flour, 0.1 units of bread, 0.15 units of capital and 0.15 units of labor.

Multi Regional Input Output (MRIO) is a specific type of Input Output table, which includes more than one economic region. Those regions could have similar or different sectors. Such table therefore provides two types of sub-tables. Diagonal tables contain the basic regional input output tables. Other tables account for international trade, illustrating flows from each sector of one country to each other sector of another country. The flows of all sector-to-sector trades sum up to form the total exports from the first country to the second. By

		Country A		Country B	
		Farms	Bakeries	Farms	Bakeries
Country A	Farms				
	Bakeries				
Country B	Farms				
	Bakeries				

Illustration 2 - Example of a MRIO table.

Blue sub-tables contain national input output accounting.

Green table represents exports from country A to country B.

Orange table represents imports from country B to country A.

¹ There is equivalence in the IO framework between both concepts. The value produced by a company equals to its total output, which arithmetically corresponds to its inputs. Those are made of intermediate and primary inputs. The firsts come from the other sectors of the economy and keep their value through the transformation process. The second correspond to the *value-added* by the company through the production process.

expanding such mapping to every country, one could then produce an interconnected overview of the world economy.

The division of the economy (or *sectorial disaggregation*) is a key characteristic of an Input Output tables. Such model assumes that each sector produces only one single output, perfectly substitutable with every other output from the same sector. The narrower those sectors are, the closer from truth this assumption becomes. However, extended disaggregation is costly, both in terms of numerical complexity and data collection. The Supply Use framework (SUT) offers some extended possibilities for goods distinction, by distinguishing industries from actual goods. In that alternative, industries do not directly use outputs from other industries, but goods previously produced by formers industries. This framework provides enhanced flexibility, because one industry could produce more than one good, and more than one industry could produce the same good. Practically speaking, Supply Use Tables (SUTs) have the same properties as IO tables. They only add numerous scarce matrices, because industries do not exchange trade directly with industry, and commodities do not interact with each other. Subsequent developments therefore apply indifferently to both frameworks.

In a nutshell, IO tables model an economy in terms of interrelations between sectors. Such model therefore highlights interdependencies between sectors: the production of each output requires the association of several intermediate goods (which are output from other sectors), along with primary inputs. It underlines that production for the final consumption is just a fraction of the total economic output: a substantial part of the production goes into several intermediate processes, before turning into a final consumption good. Using IO table, one could trace back all primary inputs required to produce a unit of final consumption. For the sake of simplicity, we illustrate this main application of IO tables using the basic example provided in Illustration I. Following results hold for any kind of IO table.

2) Notations

Let T be the trade matrix for intermediate goods:

$$T = \begin{pmatrix} 20 & 120 \\ 32 & 20 \end{pmatrix}$$

We also define matrices V and Y_0 , respectively for primary inputs and final demand. Here:

$$Y_0 = \begin{pmatrix} 20 \\ 148 \end{pmatrix}$$

$$V = \begin{pmatrix} 68 & 30 \\ 40 & 30 \end{pmatrix}$$

We define the gross output array X_0 , which sums up all outputs for each sector:

$$X_0 = Ti + Y_0 = i'T + i'V \quad (1)$$

Where subscript i is a column-summing operator (a column array of ones). Here, we have:

$$X_0 = \begin{pmatrix} 160 \\ 200 \end{pmatrix}$$

We then define the technical matrix A , from the trade matrix and the gross output:

$$A = T\widehat{X_0}^{-1} \quad (2)$$

Where the cap stands for the diagonalization operator. In our example, the technical matrix is:

$$A = \begin{pmatrix} 0.125 & 0.6 \\ 0.2 & 0.1 \end{pmatrix}$$

Conversely, we define the normalized value-added array v , embodying the quantity of labor and capital required to produce one unit of each output.

$$v = \widehat{VX_0}^{-1}$$

$$v = \begin{pmatrix} 0.425 & 0.15 \\ 0.25 & 0.15 \end{pmatrix}$$

3) Iterative approach

Let's investigate the total inputs required from the economy to eventually consume one unit of bread. We therefore define y_1 the final demand array as it:

$$y_1 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

We define the array X of wheat and bread required to meet the final demand y_1 . Of course, the consumption of one unit of bread directly requires one unit of bread. We note it $x_0 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$. As described in Illustration 1A, the production process of this unit of bread requires itself 0.6 units of wheat and 0.1 units of bread, that we note x_1 . In mathematical terms:

$$x_1 = Ay_1 \quad (3)$$

$$x_1 = \begin{pmatrix} 0.6 \\ 0.1 \end{pmatrix}$$

However, those quantities of wheat and bread do not exist *ex ante*: resources were to be used to produce x_1 . Since the production of one unit of wheat requires 0.125 units of wheat and 0.2 units of bread, the production of the 0.6 units of wheat requires 0.075 units of wheat and 0.12 units of bread. Similarly, the 0.1 units of bread require 0.06 units of wheat and 0.01 units of bread. In total, the production of the inputs to produce one unit of bread requires 0.0135 units of wheat and 0.13 units of bread. Those are *second rank* inputs, and we note them x_2 . We could compute them similarly to the computation of x_1 :

$$x_2 = Ax_1$$

And by introducing the definition of x_1 :

$$x_2 = A(Ay_0)$$

$$x_2 = A^2y_0 \quad (4)$$

We could push this reasoning one step further by computing the input of *rank 3*, required to process the *rank 2* inputs:

$$x_3 = A^3y_0$$

$$x_3 = \begin{pmatrix} 0.095 \\ 0.04 \end{pmatrix}$$

And so on until *rank n*, with the iterative formula:

$$x_n = A^n y_0 \quad (5)$$

This formula provides the array of inputs involved at *rank n* x_n to produce the final demand array y_0 . By enumerating the ranks of input, we go upward the supply chain of y_0 : each rank pictures

the ingredients required for the following rank of process to happen. Rank 1 gathers all materials required for the production process of y_0 . Rank 2 represents the inputs required for rank 1, and so on... In the end, the total input required to produce y_0 and all the ranks of inputs is the sum of the input requirements of each rank of process:

$$X = \sum_{i=0}^{\infty} x_i$$

Or, using equation (5):

$$X = \sum_{i=0}^{\infty} A^i y_0 \quad (6)$$

This demonstration holds true for any IO table and final demand array Y . By using the convention $A^0 = I$, where I is the identity matrix, the generalization of equation (5) gives, for any technical matrix A and any final demand array Y , the total input requirement X :

$$X = \sum_{i=0}^{\infty} A^i Y \quad (I)$$

This formula provides all the direct and indirect inputs required to produce any final output Y . In that perspective, our final unit of bread is the result of infinite rounds of processing, with each one feeding the next step of the processes until the production of the final output. This result includes every single step of this process and provides the total quantities of bread and wheat that were to be produced to ultimately consume one unit of bread. The mathematical convergence is ensured by the fact that all coefficients of A are natural integer strictly inferior to 1. Therefore, each power of the technical matrix has smaller coefficients, meaning that the higher we look at the supply chain, the smaller the required quantities are (at least in monetary units). While this method gives in theory access to all inputs involved in the production of any final products, it requires infinite numbers of computations. Hopefully, the Leontief inversion provides an elegant method to circumvent this objection.

4) Leontief inverse

Let suppose the production of the final demand array Y requires the total input array X . We could write equation (I) by isolating the rank 0:

$$X = Y + \sum_{i=1}^{\infty} A^i Y \quad (7)$$

We now multiply X by the technical matrix A . By doing so, we compute the direct inputs required to compute all the inputs required by Y . In other words, it means applying the logical process described in the former section to the direct inputs of Y , beginning from the rank 1 of its supply chain instead of computing them from Y , at rank 0. Using the former mathematical description:

$$\begin{aligned} AX &= AY + \sum_{i=1}^{\infty} A^{i+1} Y \\ AX &= \sum_{i=1}^{\infty} A^i Y \end{aligned}$$

Naturally, this comes to consider the same supply chain as for Y , at the exclusion of the rank 0 (which is Y itself):

$$AX = X - Y$$

Basic manipulations give then the following equation:

$$X = (I - A)^{-1}Y$$

From which we define the Leontief inverse L :

$$L = (I - A)^{-1} \quad (II)$$

Leading to the final equation of this demonstration:

$$X = LY \quad (III)$$

Equation (III) is an equivalent form of equation (I)¹. It gives by a unique operation the array of inputs X required to proceed all the steps to produce Y . Coming back to our previous example, we have the following Leontief inverse:

$$L = \begin{pmatrix} 0.875 & -0.6 \\ -0.2 & 0.9 \end{pmatrix}^{-1}$$

$$L = \begin{pmatrix} 1.348 & 0.899 \\ 0.300 & 1.311 \end{pmatrix}$$

Which directly gives us the global inputs required to produce one unit of bread:

$$X = \begin{pmatrix} 1.348 & 0.899 \\ 0.300 & 1.311 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

$$X = \begin{pmatrix} 0.899 \\ 1.311 \end{pmatrix}$$

In our fictional economy, the production of one unit of bread requires the total production of 0.899 units of wheat and 1.311 units of bread.

5) Application to primary input analysis

A further step is to translate those input requirements in terms of primary inputs. Array v offers the normalized amount of capital and labor required to produce one unit of any commodity. Let U be the work and capital required to satisfy the final demand Y . We have consequently:

$$U = vX \quad (8)$$

Or, equivalently:

$$U = vLY \quad (IV)$$

By applying this formula to the ongoing example:

$$U = \begin{pmatrix} 0.579 \\ 0.421 \end{pmatrix}$$

The production of unit of bread ultimately requires 0.579 units of capital and 0.421 units of labor. Note that this numbers are higher than what directly requires the process to create one unit of

¹ Highlighting the classic linear algebra result that $(I - A)^{-1} = \sum_{i=0}^{\infty} A^i$, as long as A has all eigenvalues strictly smaller than 1.

bread (0.15 of both). Indeed, those last numbers do not cover the previous step of the supply chain. The production of one unit of bread requires various process before the single transformation of wheat into bread. The supplementary need for primary inputs allows going through those processes.

6) Structural decomposition

Previous equations give an aggregated overview of the inputs required by the satisfaction of any final demand array Y . Let n be the number of sectors in the economy and l the number of primary inputs. Former reasoning provides the required inputs X as a column array of dimension $(n, 1)$, and required primary inputs as a line array of dimension $(1, l)$. However, such formalism does not allow:

- Accounting for the weight of the elements that compose the final demand array Y in the input requirement X .
- Accounting for the respective contribution of the inputs X to the primary inputs requirements U

Small modifications of equations (III) and (IV) offer a more detailed insights of the results:

$$\bar{X} = L\hat{Y} \quad (V)$$

Here, the upper score highlights that the primary input X now comes as a matrix instead of an array, and the cap indicates the diagonalization of a matrix. Equation (V) provides all primary inputs of the final demand array in a two-dimension table, where each column is filled with the input required to produce each type of commodity forming the final demand Y . In our example:

$$\bar{X} = \begin{pmatrix} 0 & 0.899 \\ 0 & 1.311 \end{pmatrix}$$

In that case our final demand array contains only one commodity, bread, in second position. The first column therefore only contains zeroes: the final consumption of zero units of wheat requires 0 unit of both wheat and bread. Conversely, the consumption of one unit of bread requires to produce 0.899 units of wheat and 1.311 units of bread. Row-wise sum provides the total input requirement of our final demand basket: 0.899 units of wheat and 1.311 units of bread.

There is then two main ways of expanding equation (IV):

$$\bar{U}_1 = vL\hat{Y} = v\bar{X} \quad (VI)$$

$$\bar{U}_2 = v\hat{L}Y = v\hat{X} \quad (VII)$$

In equation (VI), the diagonalization applies to the initial final demand array. So-computed matrix therefore distinguishes the primary inputs requirements of the different elements of the final demand array Y . It is therefore the direct translation of equation (V) in primary inputs:

$$\bar{U}_1 = \begin{pmatrix} 0 & 0.579 \\ 0 & 0.421 \end{pmatrix}$$

Each line stands for a kind of primary input: capital in the first row, labor in the second row. Each column represents a distinct element of the final demand: wheat in the first column, and bread in the second one. For instance, 0.579 is the total amount of capital required for the final consumption of one unit of bread.

In Equation (VII) the diagonalization applies to the intermediate inputs. This small variation triggers a sensibly different result: the result does not distinguish the allocation of primary inputs among the items of the final demand array, but among the intermediate inputs:

$$\bar{U}_2 = \begin{pmatrix} 0.382 & 0.197 \\ 0.225 & 0.197 \end{pmatrix}$$

In this case, the coefficients give insights on the final demand as a whole and do not allow distinguishing to which item of the final demand they relate. As in the last display, the first row provides capital expenses, while the second row gives the cost in labor. The first column gives the primary inputs used for producing wheat at every stage of the supply chain, while the second column provides all primary inputs used to produce bread. Contrarily to the previous display, this alternative does not allow to discriminate which item of the final demand ultimately required the mobilization of the primary inputs. However, it allows identifying how much primary inputs were mobilized in every sector of the economy to satisfy the final demand. Finally, summing this matrix over its rows gives the transposed version of the primary input array:

$$U^T = \begin{pmatrix} 0.579 \\ 0.421 \end{pmatrix}$$

Ultimately, there is another way of expanding U , when it has only one dimension (therefore when there is only one type of primary inputs):

$$\bar{U}_3 = \hat{v}L\hat{Y} \quad (\text{VIII})$$

In that last case, we diagonalize both the final demand array and the normalized primary input array. The consequence is that we distinguish both the individual items in the final demand array, and their implication on indirect inputs, measured in primary input unit. More precisely, $\hat{v}L$ gives a translation of the Leontief matrix in terms of primary input. In our previous example:

$$\hat{v}L = \begin{pmatrix} 0.337 & 0.225 \\ 0.045 & 0.197 \end{pmatrix}$$

This matrix is in the same format as the Leontief matrix: each row and each column represent a sector, and each coefficient is the total requirement of one sector over another. However, this matrix is not in monetary units anymore, but in labor units. Consequently, we read that the production of one unit of wheat for final consumption requires in total 0.337 units of farming labor and 0.045 units of baking labor. If we apply our previous final demand array, we find the following matrix:

$$\bar{U}_3 = \begin{pmatrix} 0 & 0.225 \\ 0 & 0.197 \end{pmatrix}$$

Such alternative allows displaying both the role of each item of the final demand in the global inputs' requirements, but also the repartition of primary inputs throughout the economic sectors. Note that:

- summing row-wise this matrix gives the transposed version of the labor column of $\bar{U}_2$
- summing column-wise gives the labor row of matrix $\bar{U}_1$

This last alternative version of U gives the more detailed insights on the repartition of primary inputs between both final demand and intermediate goods. However, it is only applicable for a unique primary input at a time.

7) Limits and drawbacks

IO analysis offers a consistent and functional view of the functioning of an economy. However, this approach is not a silver bullet. Several characteristics and implementation drawbacks limit its precision.

a) Sectorial aggregation

IO tables offer a simplified overview of an economy. They rely on the implicit assumption that every unit flowing out an economic sector is perfectly substitutable to any other unit from the same sector. In our small example, we therefore pretend that each piece of bread is exactly like any other. This could be wrong, for instance if bakers produce both white bread and whole bread. Our previous table does not allow capturing the specific requirements of the different types of bread. Conversely, the technique offers a sort of averaged result over the different types of commodities included in one sector. The only way to circumvent this problem is to disaggregate the baker sector into different subsectors, corresponding more precisely to the different commodities. However, higher disaggregation requires an exponential increase of the data table size. This is challenging both in terms of data collection, and in numerical complexity.

b) Data Collection

Be n the number of sectors in the economy: the constitution of an IO table requires n^2 data points. Building big data tables is therefore highly time-consuming. The more disaggregated the table are, the tougher it is to gather enough information. Raw data comes from industry surveys, which are often confidential. Many countries provide national IO tables, but their sectorial precision may be inadequate. The construction of Multi-Regional IO tables (MRIO) also requires the use of international trade databases. Those trigger additional challenges, including inconsistencies from one to another (Lenzen et al. 2012). Due to all those challenges, the constitution of IO tables often requires extrapolations and data processing, in order to fill all data points and to balance the overall table. The consequence is that data used in IO tables does not provide an exact representation of the world economy and may sometimes substantially diverge from reality. This is more the case for small sectors, for which few information is available. Additionally, the amount of work required to build the table implies that data tables for a given year are available with a substantial delay, ranging often from three to ten years.

c) A static class of models

IO tables provide a temporally aggregated view of an economy, because it considers the commercial transactions occurring over a time frame; most often on a year. This feature hinders both the variations at smaller time scales, and dynamic phenomenon. The first aspect may concern activity variation from month to month, which could alter the relative prices of commodities. For instance, the price of wheat is not the same right after the mowing and at the end of the winter. Similarly, activities running over multiple years are imperfectly captured by such class of model. Some projects indeed go through multiple phases, that will require different input and output mix over years. As a consequence, the input composition of a sector may vary over time, and the representation of a single year may not offer the full picture. For instance, IO tables poorly capture long term investments. A company that build a new building will for instance show large amounts of inputs from the construction sector for the years surrounding its construction, that will vanish once the construction is over. Consequently, IO tables poorly reflect the actualization of long-term investments over time, and so on for their primary inputs' requirements.

d) Price composition

IO tables should display the *purest* prices possible, i.e. excluding all taxes and intermediary margins, such as transportation or trade fees. Indeed, those layers on basic prices do not convey any change in the material flow. In our previous example, we could imagine that some Value-added Tax applies on goods for final consumers. However, even if their public price double compared to the price paid by the farmers and the bakers, each unit of bread still contains the same amount of primary inputs. Using the public price would then lead to unduly double the proportion of bread sold to final

consumers. That is also the case if the price of bread includes its transportation: even if farmers pay their bread 10% higher than bakers because of the distance between the bakery and the farm, both units contain the same inputs¹. Therefore, relevant IO analysis requires to use *basic prices*, or *producer prices*. While some layers are quite easy to deduct (VAT, duties...), this is not always the case. Anyway, the required treatment of raw data complexifies once again the constitution of such table.

8) Synthesis

IO tables provide a consistent overview of the interactions between economic actors. This macro-economic approach allows a systemic understanding of the supply chain of any final product. The Leontief inverse is a key tool to fill the gap between any final demand basket, and its indirect requirements throughout every sector of the economy. However, this method relies on a closed and consistent representation of very complex phenomenon. The implementation of Input Output Analysis therefore requires strong simplification and captures only a global picture of the actual situation.

¹ Accounting for the transport in our example would require the addition of a third sector to our economy, which would sum up all transportation fees not considered in basic prices. However, such solution would account globally for the transportation requirement for each sector, but not for specific transaction. The only way to do so is building a twin IO table for transportation. Some providers display indeed IO tables with different layers (transportation, trade margins, taxes, subsidies...).

I.B. Environmental Extension

Environmental Extension of Input Output Analysis (EEIO) relies on the basis introduced in the last section. Following pieces explore diverse aspects of this extension. The first part offers an overview of the related techniques. The second illustrates the specific challenges of its implementation, by extending an existing MRIOT with environmental data. The last part is a direct application of EEIO analysis to assess the Greenhouse Gases component in international trade for some major countries.

1) Presentation

EEIO is a way to apply IO analysis techniques to social and environmental issues. The basic idea is to pass the economic decomposition of the final demand array through a supplementary layer, in order to get the externalities of some consumption patterns on diverse indicators. We first shortly introduce the math of this technique. The following paragraph precise the underlying limits and assumption of such approach. We then introduce the state-of-art of the application cases f those techniques.

a) Theory

The extension of IO to socio-environmental analysis follows the exact same logic than the extension to primary inputs introduced in section I.A.1. The conceptual idea is to consider some social or environmental externalities as primary inputs for the production. In that perspective, we introduce the satellite matrix Q . This matrix is of dimension (l, n) , where l is the number of socio-environmental indicators, and n the number of sectors in the economy. It associates to every sector of the table its externalities on some indicator. Those indicators are often in a different unit compared to the rest of the table, and do not account for balancing the table. For instance, let's extend our previous example with land use. We now suppose that farms occupy a total of 180 ha, while bakeries take 20 ha. Thus, the following satellite matrix:

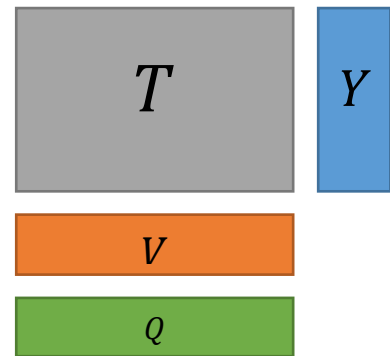


Illustration 3 – Organization of an EEIO table.

T : Intermediate table (or trade table)

V : Value-added (or primary inputs)

Q : Satellite matrix

Y : Final demand

$$Q = \begin{pmatrix} 180 & 8 \end{pmatrix}$$

Like for the value-added table, we define a normalized satellite matrix q , with x_0 the gross output:

$$q = \widehat{QX_0}^{-1}$$

$$q = \begin{pmatrix} 1 & 0.04 \end{pmatrix}$$

The production process for each unit of wheat therefore uses one hectare, and each unit of processed bread requires 0.01 hectares. Let's consider once again a final demand Y , consisting in one unit of bread. In the same manner we previously translated indirect inputs into primary inputs, we could analyze it in terms of total required surface S :

$$S = qLY \quad (IX)$$

$$S = 0.951$$

Or, in expanded version as in equation (VII):

$$\bar{S}_2 = qLY \quad (X)$$

$$\bar{S}_2 = \begin{pmatrix} 0.899 & 0.052 \end{pmatrix}$$

The final consumption of one unit of bread therefore requires the exploitation of 0.951 hectares of land. The baking process requires only 0.052 ha of land occupation. Most of the required space serves to produce wheat flour.

b) Basic applications

EEIO allows tracing back socio-environmental externalities to the final consumer caused them. Such approach, called Consumption Based Accounting (CBA), is a cornerstone of environmental policy-making. Indeed, it enables identifying not only the sectors releasing negative externalities, but also the behaviors that causes them. By offering consistent and exhaustive insights and global economic systems, EEIO is a major tool for footprint computation. In the context of increased attention on GHG emissions, carbon dioxide (CO₂) emissions is the leading topic on this field (T. Wiedmann 2009; Keiichiro Kanemoto et al. 2012; Alsamawi et al. 2014; Akizu-Gardoki et al. 2018; Global Carbon Project 2018; Jiborn et al. 2018). However, the flexibility of the method makes it open to many indicators, including local pollutants (K. Kanemoto et al. 2014), materials consumption (T. O. Wiedmann et al. 2015), water use (Lenzen et al. 2013), fertilizer use (Potter et al. 2010), biodiversity threats (Moran and Kanemoto 2017) and even social impacts (Benoit-Norris, Cavan, and Norris 2012). Application of such method at global scale highlight high scale externalities trade from developing countries to the richest ones, by the mean of international trade in goods and services (T. Wiedmann and Lenzen 2018). A substantial portion of developing countries' environmental externalities therefore serve the consumption of developed countries.

c) Resources

All EEIO applications require the coordination of an economic database with some socio-environmental data. Thanks to tremendous progress in data collection and treatment, last decade saw the emergence of several EEIO tables at world scale. The differences between those tables issue from different characteristics: number of regions, data sources, processing, available social and environmental indicators... Following table provides a small summary of the characteristics of the main tables available at the world scale. (T. Wiedmann et al. 2011) provides a more complete study on the topic, however a bit outdated.

Name	Latest available year	Number of countries	Number of sectors	Specificity
EORA	2015	190	From 26 to 300+	Exists in two versions, with different level of sectorial disaggregation
EORA 26	2015	190	26	Simplified version of EORA
GTAP	2014	121	65	
EXIOBASE	2011	44 + 5 Regions	200	Built for EEIO

Table 1 - Summary of the characteristics of the major EEIO tables at the world scale

d) Assumptions and limits

The extension of economic IO to socio-environmental issues relies on a few limits and assumptions. First, by leaning on economic IO tables, EEIO suffers from the same flaws than the previous ones. The extension is even more sensible to the broadness of sectorial aggregation since socio-environmental externalities may vary greatly from apparently close processes. A classic illustration is greenhouse gases related to electricity generation. The production of the same quantity of energy may proceed for various processes, with very different associated emissions. A table picturing a single energy production sector will miss the distinction between those different alternatives. A related issue concerns the equivalence between monetary and physical flows. The allocation of environmental externalities to final consumers follows the repartition of the formers. However, emissions associated with a particular use are not always proportionate to their economic

weight. For instance, electricity providers may charge differently households and industrial consumers, leading to an unfair allocation of externalities between both categories. A further challenge with EEIO is the requirement of exhaustivity. While most raw accounting data is available at local scale, social and environmental externalities know no borders. A large part of externalities occurs by the mean of international trade. A sound assessment of supply chain's externalities therefore requires a precise mapping of multiple regions of the world, requiring large databases. Eventually, the mapping of environmental data on economic sectors plays a crucial role on the correctness of EEIO. Upcoming Section provides a more precise insight on this last issue.

2) Integrating environmental data to economic tables: merging PRIMAP and EORA

Creating an economically relevant MRIOT requires dealing with multiple data sources, often contradictory and incomplete. Still, formatting of this information is most of the time consistent with MRIO's classifications and requires mostly aggregation or disaggregation (depending of the level of detail of both). The main challenges of the construction of a purely economic MRIOT are dealing with inconsistency and unbalances between different sources and extrapolating missing data. Integrating social or environmental indicators into a MRIOT raises different concerns, as exemplified with GHG emissions. This section reports the integration of eight non-CO2 GHGs and GHGs aggregates¹ and one CO2 indicator from an environmental database (PRIMAP-hist) to an economic MRIOT (EORA 26) By introducing a standard method for such purpose, I highlight the choices and balances the modeler must arbitrate and the consequences in terms of reliability.

EORA 26 is a widely used academic MRIOT covering 189 countries, with a high sectorial disaggregation (Lenzen et al. 2012). Being easily accessible, it includes a set of environmental indicators and supported numerous analysis of the propagation of diverse environmental externalities along supply chains, for instance water (Lenzen et al. 2013), material (T. O. Wiedmann et al. 2015), or carbon (Kander et al. 2015). However, EORA does not currently include any consistent indicator for non-CO2 GHG emissions. Contrary to economic data, the reporting of GHG emissions relies on worldly acknowledged guidelines (Eggleston et al. 2006), allowing the existence of consistent databases at world level. PRIMAP-Hist gathers multiple data sources a unique consistent set (Gütschow, Jeffery, and Gieseke 2019), available from 1850 to 2016. However, PRIMAP-hist provides too aggregated data, which does not allow a precise mapping of the emissions between economic sector. We overcome this challenge by adding a supplementary step to disaggregate PRIMAP-hist using nationally reported data (PRIMAP-crf).

This work took place at the Potsdam Institute for Climate Impact Research (PIK), with the collaboration of Johannes Gütschow². The following content is under process for further academic publication. The first part introduces both databases. The second gives an outlook on the matching between PRIMAP and EORA's matching. The third part offers a method to deal with incomplete information and inconsistencies, while the fourth and last assesses the quality of the so-computed allocation.

a) Mapping EORA Sectors with 2006 IPCC Reporting Guidelines

IPPC's guidelines released in 2006 defines more than one hundred sources of gas emissions, ordered in multiple tiers. The first step consists in selecting the relevant degree of disaggregation for each category and associating it with an EORA sector. The following table offers an overview of the

¹ CH4, FGASES, HFCS, KYOTOGHG, N2O, NF3, PFCS, SF6

² Johannes.guetschow@pik-potsdam.de

matching between economic sectors and emissions categories. Annex I provides the full discussion supporting this repartition.

IPCC Classification			EORA Sector
0	Total emissions		
	Energy		
	<i>Fuel Combustion Activities (Sectoral Approach)</i>		
1	1A	Energy Industries	
		1A1A Main Activity Electricity and Heat Production	Electricity, Gas and Water
		1A1B Petroleum Refining	Petroleum, Chemical and Non-Metallic Mineral Products
		1A1C Manufacture of Solid Fuels and Other Energy Industries	Petroleum, Chemical and Non-Metallic Mineral Products
		Manufacturing Industries and Construction	
		1A2A Iron and Steel	Metal Products
		1A2B Non-Ferrous Metals	Metal Products
		1A2C Chemicals	Petroleum, Chemical and Non-Metallic Mineral Products
		1A2D Pulp, Paper and Print	Wood and Paper
		1A2E Food Processing, Beverages and Tobacco	Food & Beverages
		1A2F Non-Metallic Minerals	Petroleum, Chemical and Non-Metallic Mineral Products
		1A2G Transport Equipment	Transport Equipment
		1A2H Machinery	Electrical and Machinery
		1A2I Mining and Quarrying	Mining and Quarrying
		1A2J Wood and Wood Products	Wood and Paper
		1A2K Construction	Construction
		1A2L Textile and Leather	Textiles and Wearing Apparel
		1A2M Other	
		1A3 Transport	Transport
		Other Sectors (Fuel Combustion)	
		1A4A Commercial/Institutional	Petroleum, Chemical and Non-Metallic Mineral Products
		1A4B Residential	Petroleum, Chemical and Non-Metallic Mineral Products
		1A4C Agriculture/Forestry/Fishing/Fish Farms	
		1A4C1 Agriculture, Forestry	Agriculture
		1A4C2 Fishing	Fishing
		1A4C3 Fish Farms	Fishing
		1A5 Non-Specified	Petroleum, Chemical and Non-Metallic Mineral Products
		<i>Fugitive Emissions from Fuels</i>	
	1B	1B1 Solid Fuels	Mining and Quarrying
		1B2 Oil and Natural Gas	Petroleum, Chemical and Non-Metallic Mineral Products
		1B3 Other emissions from Energy Production	Petroleum, Chemical and Non-Metallic Mineral Products
2	Industrial Processes and Product Use		
	2A	Mineral Industry	Petroleum, Chemical and Non-Metallic Mineral Products

3	2B	Chemical Industry	Petroleum, Chemical and Non-Metallic Mineral Products
	2C	Metal Industry	Metal Products
	2D	Non-Energy Products from Fuels and Solvent Use	Petroleum, Chemical and Non-Metallic Mineral Products
	2E	Electronics Industry	Electrical and Machinery
	2F	Product uses as Substitutes for Ozone Depleting Substances	Electrical and Machinery
	2G	Other Product Manufacture and Use	Electrical and Machinery
	2H	Other	
		2H1 Pulp and Paper Industry	Wood and Paper
		2H2 Food and Beverages Industry	Food & Beverages
		2H3 Other (please specify)	
	Agriculture, Forestry and Other Land Use		
	3A	Livestock	Agriculture
	3B	Land	
	3C	Aggregate sources and non-CO2 emissions sources on land	
	3D	Other	
4	Waste		
5	Other		
MOEL	Total CO2 equivalent emissions, without land use, land-use change and forestry, table10s1/s2/s6		
MAGELV	Total Agriculture, CRF table 3		Agriculture

Table 2 - Mapping between 2006 IPCC Guidelines and EORA Sectors.

Grey boxes represent unmapped sectors, because of data unavailability or aggregation issues

b) Extrapolating missing data

The mapping introduced in the last paragraph requires a high level of data disaggregation. PRIMAP-hist does unfortunately not offer such level of detail. The next paragraphs describe a method to extrapolate a detailed repartition of GHG emission using the reporting of Annex I Parties to the United Nations Framework Convention on Climate Change (Jeffery et al. 2018). The PRIMAP project releases a processed version of this reporting, under the designation PRIMAP-crf (for Common Reporting Format)¹. For the sake of clarity, we refer until now to PRIMAP-crf as crf. We also distinguish end-categories from aggregated categories. We name end-category a category that precisely matches a unique EORA sector, as in Table 2. End-categories sum up to constitute aggregated categories.

Method overview

Crf provides much more detailed description GHG emissions than PRIMAP-hist but limited to Annex I parties of the UNFCCC (thus excluding major emitters such as China, India or Brazil). Our method uses crf as a proxy to allocate aggregated emissions between subcategories. We use a two-step algorithm to account for both the technological characteristics of sectors, and the respective composition of countries' economy. The first step involves the computation of a first approximation of GHG emissions for each economic sector, using averaged carbon intensity. The second step ensures that the estimated value fits the aggregated value reported in PRIMAP-hist. This method also requires a preliminary step to evaluate the carbon intensity of each sector. Step 1 and 2 apply to each country, each year and each gas (except for KYOTOGHG and F-Gases, for which a special procedure applies). The preliminary step is common for every country. In order to ensure the clarity of the following, we voluntarily omit gas and year indices in notations.

¹ <https://www.pik-potsdam.de/paris-reality-check/PRIMAP-crf/>

Preliminary step: Definition of sectorial carbon intensity

The preliminary step relies on crf data and captures the qualitative difference between the sectors by defining an emission factor for each IPCC end-category i . Let E_c^{si} be the volume of emission for country c , IPCC category i , mapped with EORA sector s for any GHG. We also define W_c^i the output of the economic sector s in country c , extracted from the EORA database, in current US\$. For each economic sector, we define the emission factor:

$$e^i = \frac{\sum_{c, E_c^{si} \neq 0} E_c^{si}}{\sum_{c, E_c^{si} \neq 0} W_c^i} \quad (9)$$

This equation defines the carbon intensity as the amount of emissions in relation to the economic importance of a sector. It only uses countries for which crf contains data about the IPCC's category under examination. However, this emission factor is common to every country.

Step 1: A first approximation of disaggregated emissions

The first step of the algorithm uses the predefined intensity factor and economic data available in EORA to compute the virtual amount of emissions for each IPCC category not available in PRIMAP-hist. The process runs bottom up. It first computes end-categories, and gradually sum them up to obtain temporary aggregated categories. The following equation applies to the computation of end categories i , where $\overline{E}_c^i$ is the estimated value:

$$\overline{E}_c^i = e^i * W_c^s \quad (10)$$

Secondly, end-categories aggregate into broader categories a :

$$\overline{E}_c^a = \sum_i \overline{E}_c^i \quad (11)$$

At this point, the algorithm built a detailed estimate of one country' emissions. Note the process only applies to sectors for which no data is available. This first approximation relies on the assumption that each country of the world shares the same emission factor, and thus the same technologies. However, the current approximation already takes in account the diversity of national economic structures, because each categories' level of emission is proportional to the economic output of the corresponding sector.

Step 2: Balancing the estimated accounting

The second step aims to reconcile the estimated accounting with PRIMAP-hist's data. Contrary to the previous step, this process goes top down, and relies on available aggregated data. For each aggregated category, we define the correction ratio α_c^s as the ratio of the sourced value E_c^a against the estimation $\overline{E}_c^a$:

$$\alpha_c^s = \frac{E_c^a}{\overline{E}_c^a} \quad (12)$$

We eventually apply this ratio to the estimated values of the subcategories of a , reconciling all subcategories to the aggregated referenced value:

$$E_c^i = \alpha_c^i * \overline{E}_c^i \quad (13)$$

The algorithm progressively goes down IPCC categories' tree, gradually correcting the estimated value. This step ensures the consistency of the former estimates with available data.

Analytically, this step adjusts each emission factor by the correction ratio α_c^S . This ratio is common to every subcategory belonging to the same category in one country. One could interpret it as a correction applied to the mean technological factor e^i , embodying the technological difference between one country and the mean value computed using crf. By combining a sectorial emission factor and a local correction ratio, our model captures differences between countries both due to economic structured and efficiency of local technologies.

Technical issues

The next paragraphs describe a few technical challenges that we encountered during the implementation of this method.

For a few emission categories, there is no data available at all (1A4C1, 1A4C2 and 1A4C3). In that very particular case, we compute the intensity factor of category 1A4C by considering the economic output as the sum of each sector associated with a subcategory of 1A4C. We then use a common intensity factor for each of the subcategories. Crf does not display the most detailed disaggregation for each country, and present some inconsistencies mainly related to rounding. We consequently apply the full process to the countries of Annex I using only crf data, to build a fully consistent starting matrix. In addition, crf reporting does not perfectly meet PRIMAP-hist data. In order to keep the results consistent, we systematically override crf aggregated data with PRIMAP-hist's equivalents, while keeping the relative repartition between subcategories as observed in crf. Therefore, some values available in crf are adjusted along the process and may differ from raw crf data.

Eventually, the KYOTOGHG4 and F-Gases aggregates undergo a different treatment, to ensure its consistency with the rest of GHG emissions data. Instead of computing it via intensity factor, we directly estimate it by using the repartition of the others GHGs, weighted by their respective Global Warming Potential Values, extracted from AR4 (Forster et al., 2007). We then apply the second step of the algorithm to ensure the consistency of the repartition with available data.

Extension to multiple data sources

Previous paragraphs describe the application of our method to two data set: a reference data set (PRIMAP-hist), and a more detailed but incomplete data set to serve as a proxy for unreferenced categories (PRIMAP-crf). Minor modifications in the process may let space for integrating more than two data sources. Following entries evoke step by step the modifications such implementation would require.

The definition of intensity factor allows extrapolating unknown data for some countries. Depending on the data available in different data sources, user may associate the definition of some sector's intensity factor with one particular data set, or with a combination of multiple data sets. A sound way to adapt the two-step iterative to process is applying it successively to each data set, ending with the most *trustful* data set. This ensures extracting the most information possessing of each data set, while keeping a consistent estimation with all available data at each iteration. A marking system allows keeping track of the successive changes on the estimation.

- When no data is available, marker vaults 0
- When data has been estimated, marker equals 1
- When data exists in a dataset, marker is 2

For each successive data source, step one estimates values for all 0 and 1 marked values. During step 2, all values covered by a known aggregated category are balanced, ensuring the global consistence of the estimate. At this point, we have a consistent estimation, which exploits all the previously integrated data sets. When considering a supplementary data set, the algorithm imports all data available in the

new dataset, overriding former estimation when necessary. Then step 1 re-initializes the estimation of uncertain values, and step 2 balances the table to make it consistent with the last information introduced. For branches where the new data set introduced no new information, the process conserves the former estimate. Branches where new information are available are reprocessed to become consistent with the new information.

c) Validation

A first overview of CO₂ emissions repartition between sectors and countries brings us a first insight on the quality of our method. For the sake of clarity, we conduct this investigation over year 2000 only. Further steps of validation will ensure the consistency of our method along time series.

Figure 1 gives us a first overview of the allocation of emissions among sectors and countries, as compared to a database already mapped in EORA 26, proceeding from the EDGAR database (Janssens-Maenhout et al. 2017).

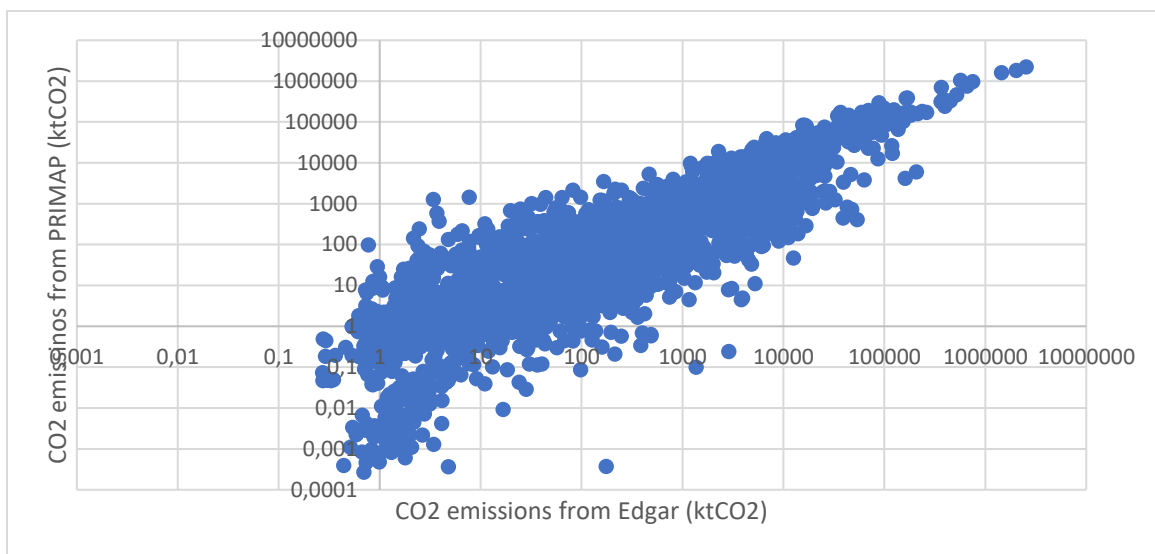


Figure 1 - Comparison of sectorial emission repartition with EORA and PRIMAP, year 2000

Each dot represents an entry for a specific sector and country, with the value reported in EORA on X-axis and the value reported in PRIMAP on Y-axis. With a perfectly equivalent match between the two measures, each dot would be aligned on the axis $Y=X$. However, our measure diverges slightly from reference measure, with a R^2 of 0.93.

I.B. Environmental Extension

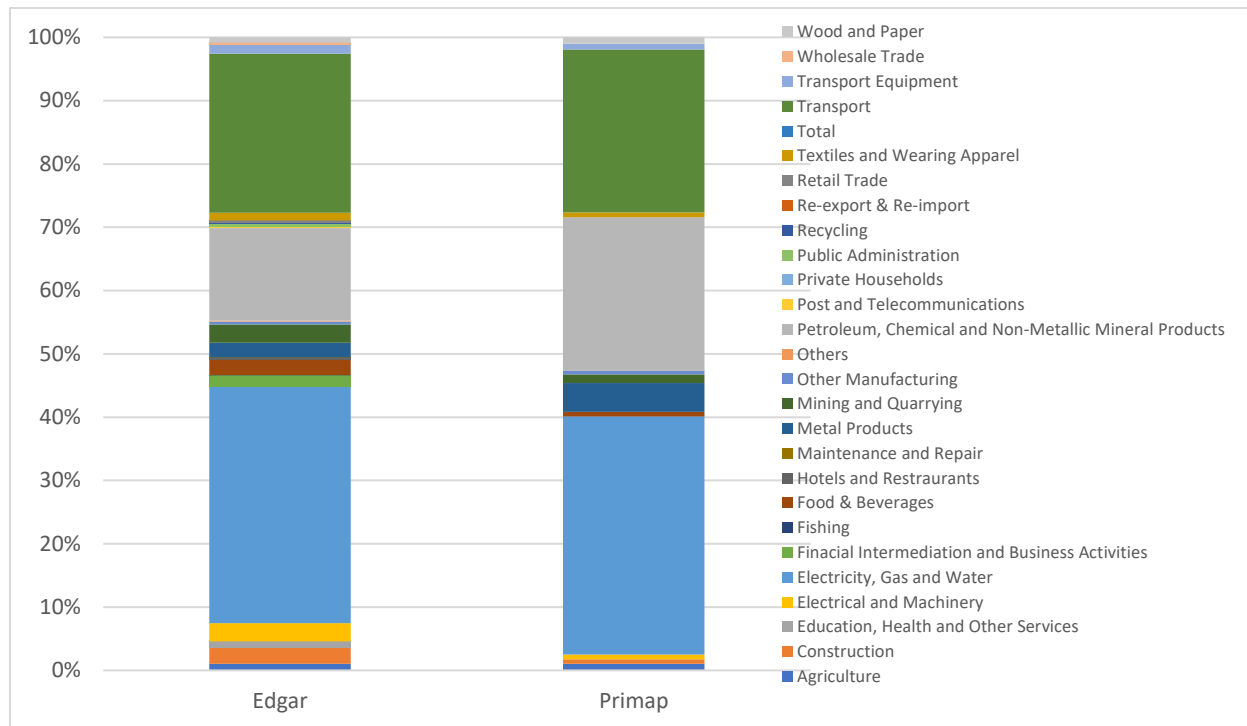


Figure 2 - Comparison of sectorial repartition of CO2 emissions between PRIMAP and Edgar

Figure 2 displays the global repartition of CO2 emissions among EORA sectors. Compared to the reference dataset, our allocation gives more importance to primary sectors (Petroleum, Chemical and Non-Metallic Mineral Products, Electricity, gas and Water), at the expense of secondary sectors (Food and Beverages, Financial Intermediation and Business Activities, Electrical and Machinery, Construction). The design of our method is consistent with those differences, because when emissions categories relate to multiple economic sectors, we link those emissions to primary sectors providing the product that causes the emissions. For instance, we assign emissions related to commercial fuel combustion (1A4A) to the petroleum sector. Edgar probably proceeds from different assumptions and assigns more emissions to secondary sectors. The sectorial allocation provided by our method is eventually slightly different from the reference dataset Edgar in EORA. However, methodological design choices explain those differences. At the same time, EEIO analysis techniques reallocate the emissions throughout sectors. Divergences in the sectorial repartition of emissions does therefore not necessarily alter further analysis.

Footprint measuring is equivalent of Consumption-Based Accounting. By allocating emissions to final uses rather than producing sector, one can identify the responsibility of consumption patterns on GHG emissions. Being the main technique of EEIO analysis, it is a proper test to assess the relevance of our allocation method. The EORA team provides a reference database for national footprint accounting¹, computed by using the Edgar databases included on EORA. By computing and comparing national footprint based on the repartition provided with our method, we could assess the holistic accuracy of our method. Note that those estimates rely on a more disaggregated table than EORA 26. Differences in results may therefore come from both differences in environmental data mapping and in economic modelling of the world economy.

¹ <https://worldmrio.com/footprints/carbon/>

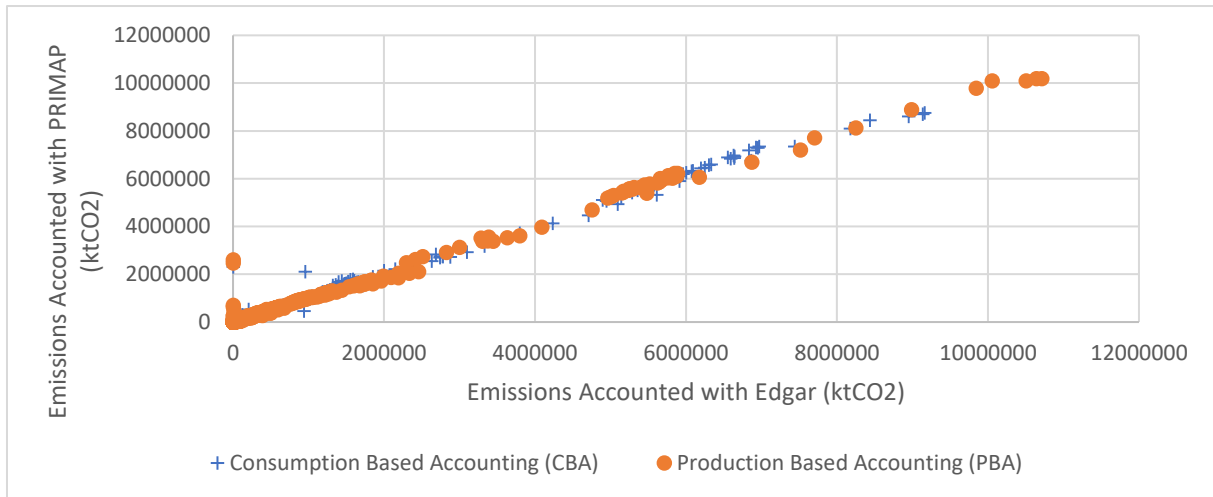


Figure 3 - Comparison of national CO2 emissions accounting using Edgar and PRIMAP

Figure 3 displays national emissions for all countries between 1990 and 2015, with reference measure on X-axis and measure obtained with our method on Y-Axis. Each orange point represents a measure for one year and on country following the Production Based Accounting (PBA) principles and each blue cross marks a measure of the Consumption Based Accounting (CBA) for one year and one country. The former measures territorial emissions of a country, while the second corresponds to the emissions associated with the final consumption of each country. Visually, both measures seem close from one to another: most dots are close to the axis $Y=X$. Statistically, both methods come with a R^2 coefficient higher than 0.99: 0.9938 for CBA and 0.9908 for PBA.

Difference between CBA and PBA represents the transfer of carbon emissions due to one country trade. This measure allows capturing more precisely the differences in emission allocation between our method and Edgar.

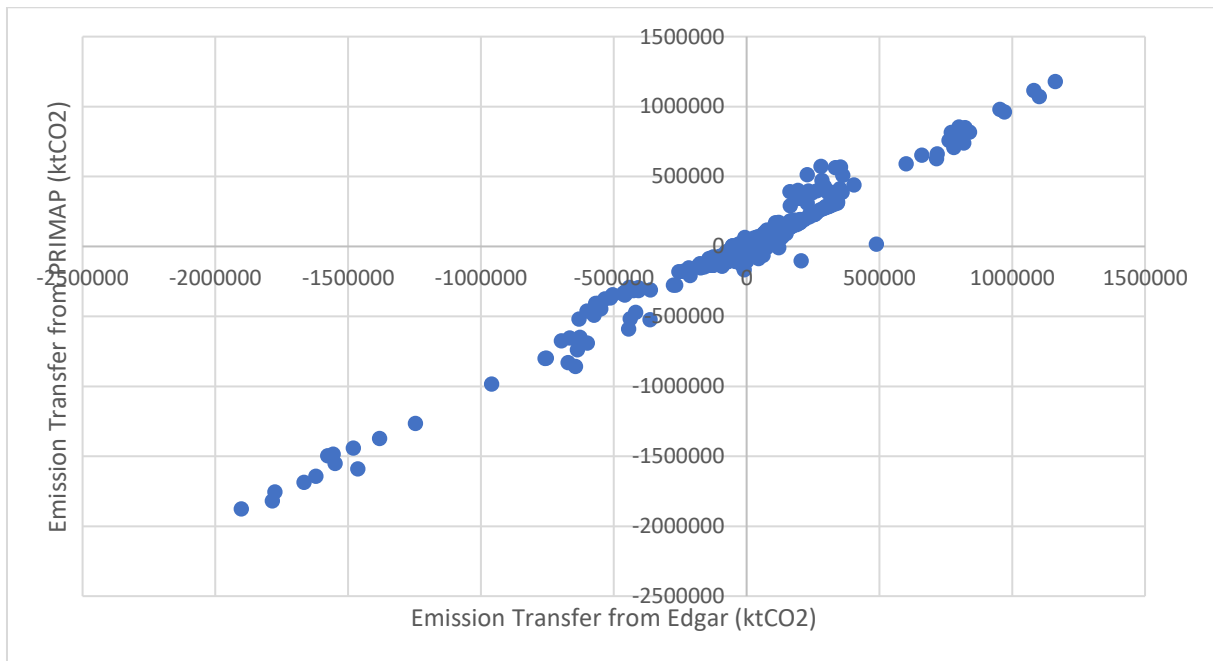


Figure 4 - Comparison of CO2 emissions transfer accounting between Edgar and PRIMAP

As illustrated by Figure 4, this measure seems also quite consistent with both databases, with a R^2 score of 0.97. Holistic measures of consistency between our allocation method and the reference

database appear to be convincing, with a strong correlation in footprint accounting for both databases. Our method is eventually in line with the existing literature.

d) Synthesis

The integration of environmental information on economic databases is a crucial challenge. The main issues lie on linking externalities categories and economic sectors, along with the scarcity and potential inconsistencies of source data. This example introduced a systematic method to overcome such challenges, using two main databases. The two-step algorithm uses the most disaggregated data available to produce a median estimate of GHG emissions for all economic sectors in the world. A top-down iterative step then scales the estimates to meet the best aggregated available measures. Such double-step method ensures taking the most information possible, accounting for both the specificity of economic sectors and between countries.

A main insight of this study case lies in the required modelling choices. Mapping IPCC's categories with economic sector is not an easy task, and different theoretical approaches provide different allocations results. However, the quantitative assessment of our results is very encouraging. Advanced analysis relying on our allocation provide results remarkably in line with reference data. This highlights that despite slight differences in methodological design, results are still interestingly reliable. A further aspect of this study is the need for compatible disaggregation levels between environmental and economic data. Mapping environmental categories on economic sectors requires the former to be assignable to precise sectors. Reciprocally, environmental data available at world level only represent broad categories so that data may be unavailable for too disaggregated data tables. This issue of compatible disaggregation is key for further development of EEIO tables and reduce the need for more precise MRIOT. Conversely, this is also a limit for the mappable indicators, since they have to be assignable with sufficient precision to economic sectors.

3) Study case: Consumption Based Accounting of global Greenhouse Gases Emissions

The last section introduced the integration of the PRIAMP-hist database into EORA. This work allows measuring footprints associated with the main categories of GHG, while this measure was only available for CO₂ emissions until now. We therefore use this new database to compare the territorial emissions five countries to their Consumption-Based Emissions. We focus on the balance between those measures, in order to identify which countries are net exporters or importers of emissions. To our knowledge, it is the first available measure of such quantity for non-CO₂ GHG emissions.

a) Method

The measure of the footprint from a country comes directly from Equation (IV), by using the final demand array corresponding to the final consumption of a country. It embodies nationally produced goods as imported goods. The result of this consumption-based account therefore integrates both the emissions required for locally produced goods as much as imported emissions. EORA provides all the required data for 2015, including final demand array for all the countries. In order to produce comparable results, results are computed by capita, using 2015 World Bank accounts ("Population, Total" 2019). GHG emissions data comes from the previously described project. For the sake of clarity, we focus on the five main categories of GHG emissions: CO₂, CH₄, N₂O, fluorinated gases (F-Gases) and the aggregated basket of gas covered by Kyoto's agreements (KYOTOGHGAR4).

b) Results

Figure 5 provides the composition of the carbon footprint for the average Chinese, French, German, British and US citizen, both in terms of territorial emissions (PBA) and footprint (CBA).



Figure 5 - Composition of GHG emissions for five major countries, in tons of CO₂ equivalents and per inhabitants. PBA stands for Production-Based Accounting and represents territorial emissions. CBA stands for Consumption-Based Accounting: it includes imported emissions and excludes the exports.

The five countries of the selection have their territorial emissions ranging from 6.8 to 20.5 ktCO₂e/capita. UK and France have the lowest level of territorial emissions, while the average American emits at least twice more GHG than any other average citizen of the panel. In each case, CO₂ emissions represent the largest contributor to GHG emissions, with between 77% and 86% of global GHG emissions. Respectively related to livestock and agriculture, methane and nitrous oxide form between 5% and 11% and between 4% and 9% of the total emissions. F-Gases form the residual emissions, ranging from 2% to 5%.

The picture is sensibly different when we look at consumption related emissions, highlighting that a substantial part of GHG emissions transit throughout international trade. China has a smaller consumption footprint compared to its production emissions: more than one ton of CO₂ equivalent emissions serve to produce goods or services that will be consumed somewhere else, which means that 13% of Chinese territorial emissions relate to other countries' consumption. Conversely, accounting for trade increases the footprint of western countries. The figure can go up to 62% for the UK, which footprint goes from 7.7 ktCO₂e/cap to 12.4 ktCO₂e/cap when accounting for trade related emissions. In France, it goes up by 42%, to 9.6 ktCO₂e. Interestingly, the difference is smaller for Germany, from 11.1 ktCO₂e/cap to 12.1 ktCO₂e/cap, representing a 9% difference.

One could notice that the accounting of international trade tends to smooth the differences between European Countries. While there is a 39% and 31% difference in territorial emissions between respectively France and Germany and England and Germany, the gap closes when accounting for trade embodied emissions. It even inverses between England and Germany, the former being 2% higher than the last. Between France and Germany, the gap tightens to a 20% difference.

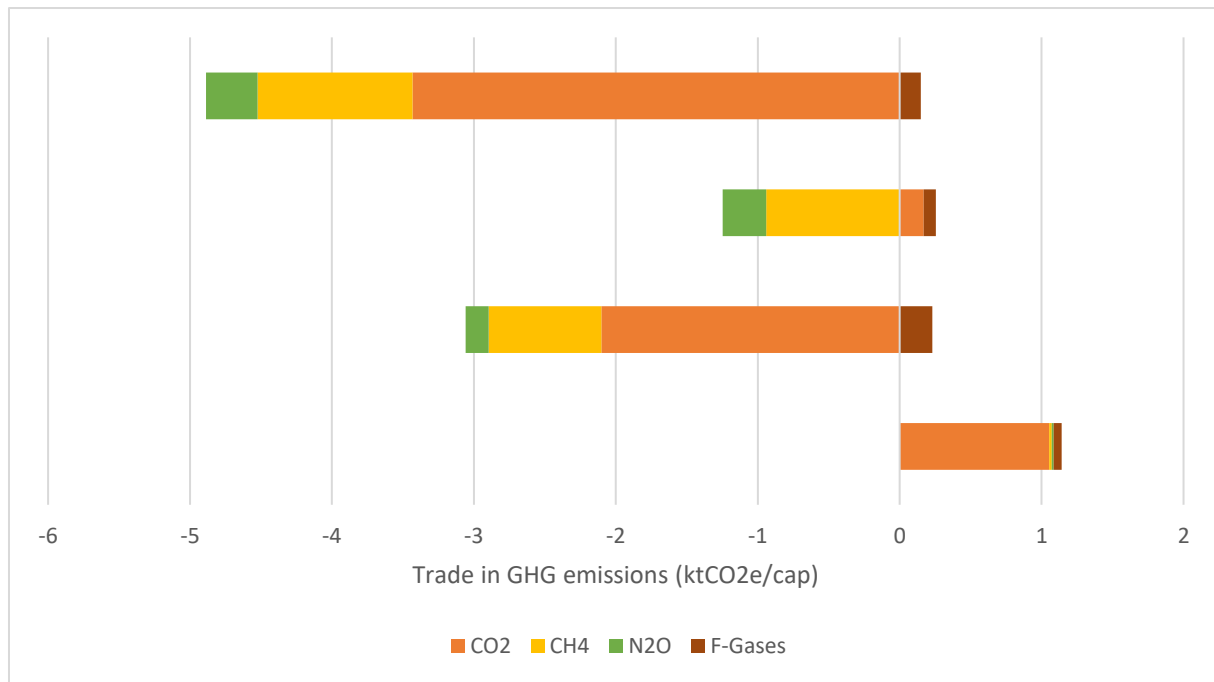


Figure 6 - Balance of GHG emissions embodied in international trade, in kilotons of CO2 equivalent per inhabitants

Figure 6 offers a deeper insight on this issue, by detailing the difference between Product-Based Accounting and Consumption Based Accounting for each country and each GHG. This figure highlights slight differences between countries. China is the only country to be net exporter of GHG emissions for every gas. Also, all countries of the panel export a substantial part of their F-Gases emissions, from 23% for China to 76% for France. Indeed, fluorinated gases are associated with advanced industrial processes, such as refrigeration or electronics. Since all the countries of the panel are technologically advanced, they produce advanced goods for other parts of the world, and a part of their F-Gases emissions go along with their exporting industry to the rest of the world. Germany is the only western country not to import CO2 emissions, with even exporting some 0.2 tons throughout national trade. The composition of the German electricity mix may partly explain this singularity, at least to France and England. By still largely relying on coal and lignite, Germany has a much more carbon-intense energy mix than its neighbors. German industries therefore export relatively more carbonized goods to its neighbors, while importing relatively low carbon-intense goods and thus reducing its consumption footprint. Eventually, western countries import methane and nitrous oxide, probably by the mean of agricultural imports.

c) Conclusion and discussion

This short study case highlights the possibility offered by EEIO for nation-wide footprint measures. This approach of course carries some drawbacks, including quantitative uncertainties. The Global Carbon Projects gives comparable measures for CO2 emissions, which may slightly differ from our estimations (Global Carbon Project 2018).

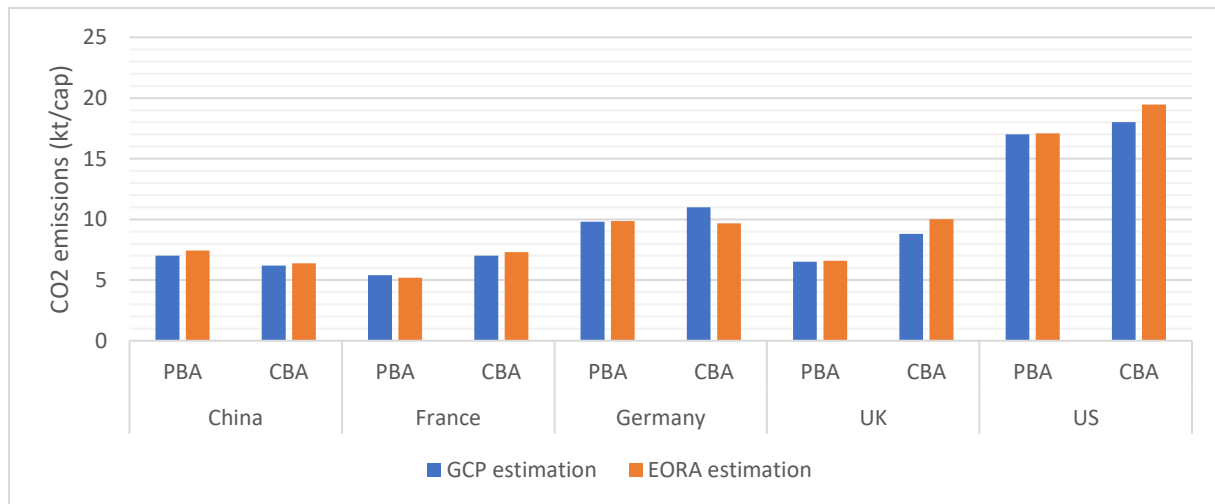


Figure 7 - Comparison of territorial and consumption footprint for CO2 emissions between GCP and EORA estimates, in 2015

Figure 7 highlights that the same method applied to different databases may provide slightly different results, those differences coming from both the environmental and the economic database. The accounting of territorial CO2 emissions is not equivalent between the two databases, ranging to 6% for Chinese emissions. This may come from different methodologies in emissions accounting and differences in data sources. Secondly, GCP relies on a different economic database than EORA. Both sectorial disaggregation and numeric values vary between the two, causing different mapping of interlinkages throughout the world economy, and therefore a different allocation of territorial emissions to exports and imports. This is obvious for Germany and US: both show limited differences in territorial emissions, but substantial differences in consumption footprint. For the case of Germany, EORA suggests that Germany is a net exporter of CO2 emissions, while GCP supports the contrary. While such difference limits the numerical certainty of the results, it appears that global results remain consistent in order a magnitude from one to another.

Ultimately, this study case calls for further analysis. When analyzing the results, assumptions on the differences in economic structures appear as convincing in theory but would require more information to be formally supported. The tool introduced in the next chapter could precisely support such analysis.

CHAPTER II: A TOOL TO ASSESS THE SOCIAL AND ENVIRONMENTAL FOOTPRINT OF SUPPLY CHAINS

The impact of human activity on Nature and local communities is a growing preoccupation in the 21st century. Repeated scientific warnings triggered the awareness of both consumers, public powers and stakeholders on the negative externalities of careless industrial activities. Consequently, pressure goes upward on private companies to account and report for their externalities. The global assessment of one activity's impacts on the environment and on local communities is therefore a burning topic. PwC France's Sustainability Performance Strategy (SPS) team has been pushing economic actors along this way since 1993, and constantly look for new ways to tackle such challenge. This chapter details the development of a new tool to address this challenge, based on EEIO principles.

The following content derives from the development report delivered to PwC France at the end of the internship period. The first part provides an overview of the tool's functionality, shortly introducing some application cases. Second part focuses on data and numerical precision and underlines its compatibility with some pre-existing methods such as Life Cycle Analysis (LCA). The third focuses on the technical implementation of the tool, highlighting the main design choices. The last part synthesizes the discussion and opens paths for further development.

II.A. Overview

The tool offers a solution to assess a broad range of externalities associated with any industrial supply chain, by implementing Equation (IV). An Excel file makes the input interface. The user defines the final demand array Y and chose the and environmental indicators Q within a list. The algorithm returns one or the other version of the matrix of externalities U . All three parameters are highly customizable, enhancing a flexible use of the tool. The execution of a standalone application proceeds on its own to all required computation and returns the results as a distinct Excel file. The tool uses two databases, with different sectorial definition. The first one, EORA 26, uses a uniform 26 sectors disaggregation. The second one, Full EORA, uses the most detailed disaggregation available for each country, ranging from 26 to more than 300 economic sectors. Both cover 190 countries. User could use simultaneously the two databases, with different parameters. The use of the EORA database relies on a permanent commercial license, which only includes 2015 for the moment and for unlimited use. However, the design of the tool could let the user to choose a different year for the database.

1) Indicators

The tool includes a total of fourteen social and environmental set of indicators, along with six economic indicators. Those indicators cover various dimensions, such as GHG emissions, material and water use, nitrogen disposal, employment, energy mix or the repartition of value-added. User could select multiple indicators at once. The algorithm uses the selected indicators for both models. However, the quality of the indicators varies between the models. Some indicators are to use preferably with one model over the other. Note that the indicators extracted from PRIMAP (see Section I.B.2) are not available in Full EORA. Section II.B. gives more details on this issue.

2) Treatment modes

The *Treatment mode* parameter characterizes simultaneously the composition of the final demand array Y and the type of results to compute. User defines this parameter for each model separately. There are five possibilities:

- *No Treatment* does not compute anything for the chosen model
- *Direct Impact* displays the direct externalities (rank 0 externalities) caused by one or several economic sectors manually chosen.

All three other modes compute a footprint measure, which includes both direct and indirect externalities. They differ by the input to inform.

- *National Footprint* computes the footprint associated with one country's final emissions. It includes territorial emissions and emissions embodied in imports, and excludes emissions embodied in export. This mode only requires the selection of a country in a rolling list, as illustrated in Illustration 4.

	Treatment Mode	Display
Eora 26	Manual Input	Detailed
Full_EORA	Industry Footprint	Detailed
Country For National Footprint mode		
	France	61
	France	
	French Polynesia	
	Gabon	
	Gambia	
	Gaza Strip	
	Georgia	
	Germany	
	Ghana	

Illustration 4 - Zoom in the mode selection table. The opened rolling list allows choosing the country for the "National Footprint" mode

- *Industry footprint* requires the selection of one or several industries and countries in the EORA 26 or FullEORA tabs. The algorithm then computes the global footprint of those sectors using available information on their respective economic weight from the economic database.
- *Manual input* requires the selection of countries, sectors, and the transaction volume for each of them. It then computes the footprint associated with those sectors, for the defined transaction volume. It is the most flexible input mode.

3) Input data

Input data used by the operator could correspond to different economic objects, depending on the scope of the study. For investigating the externalities of consumer choices, one could use expenses survey. An approach to assess the global footprint of a company is considering the activity of the company as what it sells to its customers. Consequently, the revenue made by a company on different market sectors is also a relevant output. At a product scale, the tool could also offer an outlook the externalities from the previous processes of one product's supply chain. In that perspective, user should consider the direct furnishers of the company. From their point of view, the company under study acts as a final consumer. Note also that the database uses *basic prices*, in \$US of the current year. User should therefore convert the monetary values the good currency and exclude taxes and intermediary margins from the prices.

In practice, the input file has one tab for the input data of each model. For each item the user wants to integrate in the final demand array, he must inform the country and the sector of origin of the flow, along with the monetary value associated with the transaction (except for the *Industry Footprint* mode, which does not require monetary values). User should choose the country and the sector in rolling lists. A checking system aims at easing the fulfillment of those sheets, as showed in Illustration 5.

Sector Selection	Sector Match	
Electrical energy, gas, steam and hot water	Electrical energy, gas, steam and hot wa	
Textiles and Wearing Apparel	Textiles	
Electrical and Machinery	Education services	
Computer and related services	#REF!	

Illustration 5 – Verification system for the sector selection

The user entered four sector names in the “Sector Selection” column. The first verification box (on the right) is green: there is an EORA sector (in the middle, in grey) matching the input sector (on the left, in yellow). Second and third verification boxes are orange: the input sector does not exist in EORA, but there is a close match. In the first case, the automatic matching appears correct: “Textiles” and “Textiles and Wearing apparel” are close to each other. The matching in the third line is much less satisfactory: “Electrical and Machinery” has nothing to do with “Education Services”. There is no matching at all for the fourth line. For those last two lines, user should select a more appropriate sector in the “Sector Selection” column, using the rolling menu.

4) Output

The output is an Excel file, which presents four tabs. The first displays some generic information on the computation process, such as the version of the algorithm, the time of execution... The last one repeats the input data. The second and the third display the results of the computation for the two models. The application offers different *display modes*, corresponding to alternative formatting of the output.

- The *Consumer view* display follows equation (VI). The output presents independently the impact of each item of the final demand array on each indicator.

J	A	B	C	D	E	F	G	H	I
1	Departem	Divisions	Country	Entity	Sector	Aggregated name	GHG emissions from PRIMAP - CH4 (kt)	GHG emissions from PRIMAP - CO2 (kt)	GHG emissions from PRIMAP - FGASES - AR4 (ktCO2e)
2			China	Commodi	Agriculture	- - China - Commodities - Agriculture	7.250750077	196.2031305	3.835876653
3			China	Commodi	Transport Equipment	- - China - Commodities - Transport Equipment	0.266947948	81.00241175	2.781927002
4			China	Commodi	Food & Beverages	- - China - Commodities - Food & Beverages	0.603809439	50.73502582	0.836158257
5			China	Commodi	Textiles and Wearing Apparel	- - China - Commodities - Textiles and Wearing Apparel	0.253576246	43.49107459	0.811127205
6			China	Commodi	Electricity, Gas and Water	- - China - Commodities - Electricity, Gas and Water	0.167220622	143.1109449	0.436008347
7			India	Industries	Textiles and Wearing Apparel	- - India - Industries - Textiles and Wearing Apparel	0.676330202	52.1461046	0.55974852
8			India	Industries	Agriculture	- - India - Industries - Agriculture	1.435573287	12.64770159	0.141236219
9			India	Industries	Food & Beverages	- - India - Industries - Food & Beverages	0.215299617	6.869940021	0.061654985
10			Pakistan	Industries	Textiles and Wearing Apparel	- - Pakistan - Industries - Textiles and Wearing Apparel	0.057422171	1.924703775	0.002614619
11			Pakistan	Industries	Agriculture	- - Pakistan - Industries - Agriculture	0.181389607	0.114208731	0.000220789
12			Pakistan	Industries	Transport	- - Pakistan - Industries - Transport	0.001297155	0.864056513	0.000221104
13			Pakistan	Industries	Food & Beverages	- - Pakistan - Industries - Food & Beverages	0.021864933	0.110642104	0.000184262

Illustration 6 – Example of output in consumer view. Input sectors are in rows and indicators in columns. The selected box highlights that the demand in the Indian agricultural sector causes 12.6 kt of CO2 emissions.

- The *Producer view* uses equation (VII). The output is split between the intermediate inputs required for the satisfaction of the final demand.

II.A. Overview

	A	B	C	D	E	F	G	H
1	Depart	Division	Country	Entity	Sector	Aggregated name	GHG emissions from PRIMAP - CH4 (kt)	GHG emissions from PRIMAP - CO2 (kt)
1016			China	Commodi	Agriculture	- - China - Commodities - Agriculture	7.066079105	0.277508134
1017			China	Commodi	Fishing	- - China - Commodities - Fishing	0.005895428	0.003705101
1018			China	Commodi	Mining and	- - China - Commodities - Mining and Quarrying	0.935531826	7.698089335
1019			China	Commodi	Food & Be	- - China - Commodities - Food & Beverages	0.000255558	3.761598752
1020			China	Commodi	Textiles and	- - China - Commodities - Textiles and Wearing Apparel	0.000123378	0.727485733
1021			China	Commodi	Wood and	- - China - Commodities - Wood and Paper	7.48376E-05	0.503800068
1022			China	Commodi	Petroleum	- - China - Commodities - Petroleum, Chemical and Non-Metallic Mi	0.087692252	113.5963096
1023			China	Commodi	Metal Pro	- - China - Commodities - Metal Products	0.000977908	10.23142847
1024			China	Commodi	Electrical	- - China - Commodities - Electrical and Machinery	1.95027E-06	0.176325278
1025			China	Commodi	Transport	- - China - Commodities - Transport Equipment	0.000291998	1.008248127
1026			China	Commodi	Other Man	- - China - Commodities - Other Manufacturing	0	0
1027			China	Commodi	Recycling	- - China - Commodities - Recycling	0	0
1028			China	Commodi	Electricity	- - China - Commodities - Electricity, Gas and Water	0.007055412	284.3993382
1029			China	Commodi	Construct	- - China - Commodities - Construction	1.71786E-06	0.011281952
1030			China	Commodi	Maintenai	- - China - Commodities - Maintenance and Repair	0	0
1031			China	Commodi	Wholesale	- - China - Commodities - Wholesale Trade	0	0

Illustration 8 – Example of output in production view. This example was generated using the same input as Illustration 6. The selected cell highlights that the production of the final demand array causes indirect emissions of 112 kt of CO2 in the Chinese Petroleum, Chemicals and Non-Metallic Minerals sector. This visualization does not distinguish the contribution of individual items of the final demand array on this value.

- The mode *By rank* is an adaptation of equation (I). It distinguishes externalities both by intermediate inputs, and by their rank in the supply chain. For technical reasons, this display mode is only available for EORA 26.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X
1	Depart	Division	Country	Entity	Sector	Aggregated name	PRIMAP - CO2 (kt) - Rank	PRIMAP - CO2 (kt) - Rank	PRIMAP - CO2 (kt) - Rank	PRIMAP - CO2 (kt) - Rank	PRIMAP - CO2 (kt) - Rank	PRIMAP - CO2 (kt) - Rank	PRIMAP - CO2 (kt) - Rank	PRIMAP - CO2 (kt) - Rank	PRIMAP - CO2 (kt) - Rank	PRIMAP - CO2 (kt) - Rank	PRIMAP - CO2 (kt) - Rank	PRIMAP - CO2 (kt) - Rank	PRIMAP - CO2 (kt) - Rank	PRIMAP - CO2 (kt) - Rank	PRIMAP - CO2 (kt) - Rank	PRIMAP - CO2 (kt) - Rank	PRIMAP - CO2 (kt) - Rank	
1016			China	Commodi	Agriculture	- China - Commodities - Agriculture	0.229866167	0.029698412	0.010500924	0.003921148	0.003521484	0.003521484	0.003521484	0.003521484	0.003521484	0.003521484	0.003521484	0.003521484	0.003521484	0.003521484	0.003521484	0.003521484	0.003521484	
1017			China	Commodi	Fishing	- China - Commodities - Fishing	0	0.000701209	0.001126959	0.000443986	0.000443986	0.000443986	0.000443986	0.000443986	0.000443986	0.000443986	0.000443986	0.000443986	0.000443986	0.000443986	0.000443986	0.000443986	0.000443986	
1018			China	Commodi	Mining and	- China - Commodities - Mining and Quarrying	0	0.930749563	2.043369944	1.643381757	3.080580071	7.698089335	7.698089335	7.698089335	7.698089335	7.698089335	7.698089335	7.698089335	7.698089335	7.698089335	7.698089335	7.698089335	7.698089335	
1019			China	Commodi	Food & Be	- China - Commodities - Food & Beverages	1.960993756	1.151906723	0.370895438	0.144183386	0.133619449	0.133619449	0.133619449	0.133619449	0.133619449	0.133619449	0.133619449	0.133619449	0.133619449	0.133619449	0.133619449	0.133619449	0.133619449	
1020			China	Commodi	Textiles ai	- China - Commodities - Textiles and Wearing Apparel	0.38915855	0.169105082	0.080879097	0.040519137	0.047823866	0.047823866	0.047823866	0.047823866	0.047823866	0.047823866	0.047823866	0.047823866	0.047823866	0.047823866	0.047823866	0.047823866	0.047823866	
1021			China	Commodi	Wood and	- China - Commodities - Wood and Paper	0	0.127703464	0.144946259	0.092434154	0.13874132	0.13874132	0.13874132	0.13874132	0.13874132	0.13874132	0.13874132	0.13874132	0.13874132	0.13874132	0.13874132	0.13874132	0.13874132	
1022			China	Commodi	Petroleum	- China - Commodities - Petroleum, Chemical and Non-Metallic Mi	0	43.81758147	27.68587781	16.76077139	25.32827675	113.5963096	113.5963096	113.5963096	113.5963096	113.5963096	113.5963096	113.5963096	113.5963096	113.5963096	113.5963096	113.5963096	113.5963096	
1023			China	Commodi	Metal Pro	- China - Commodities - Metal Products	0	1.978545592	2.264438231	1.904421837	4.084019811	10.23142847	10.23142847	10.23142847	10.23142847	10.23142847	10.23142847	10.23142847	10.23142847	10.23142847	10.23142847	10.23142847	10.23142847	
1024			China	Commodi	Electrical	- China - Commodities - Electrical and Machinery	0	0.041812272	0.042921421	0.032360361	0.059238524	0.176325278	0.176325278	0.176325278	0.176325278	0.176325278	0.176325278	0.176325278	0.176325278	0.176325278	0.176325278	0.176325278	0.176325278	
1025			China	Commodi	Transport	- China - Commodities - Transport Equipment	0.561094219	0.228773751	0.104435094	0.051465659	0.062479403	1.008248127	1.008248127	1.008248127	1.008248127	1.008248127	1.008248127	1.008248127	1.008248127	1.008248127	1.008248127	1.008248127	1.008248127	
1026			China	Commodi	Other Man	- China - Commodities - Other Manufacturing	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
1027			China	Commodi	Recycling	- China - Commodities - Recycling	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
1028			China	Commodi	Electricity	- China - Commodities - Electricity, Gas and Water	93.70319763	56.53461319	44.16223213	31.87034898	58.12894627	284.3993382	284.3993382	284.3993382	284.3993382	284.3993382	284.3993382	284.3993382	284.3993382	284.3993382	284.3993382	284.3993382	284.3993382	
1029			China	Commodi	Construct	- China - Commodities - Construction	0	0.003076032	0.003020439	0.002013298	0.003172184	0.011281952	0.011281952	0.011281952	0.011281952	0.011281952	0.011281952	0.011281952	0.011281952	0.011281952	0.011281952	0.011281952	0.011281952	
1030			China	Commodi	Maintenai	- China - Commodities - Maintenance and Repair	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
1031			China	Commodi	Wholesale	- China - Commodities - Wholesale Trade	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
1032			China	Commodi	Retail Tra	- China - Commodities - Retail Trade	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
1033			China	Commodi	Hotels and	- China - Commodities - Hotels and Restaurants	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
1034			China	Commodi	Transport	- China - Commodities - Transport	0	21.2803283	16.4082461	10.63468276	17.35495895	65.67827064	65.67827064	65.67827064	65.67827064	65.67827064	65.67827064	65.67827064	65.67827064	65.67827064	65.67827064	65.67827064	65.67827064	
1035			China	Commodi	Post and T	- China - Commodities - Post and Telecommunications	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
1036			China	Commodi	Financial In	- China - Commodities - Financial Intermediation and Business Act	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
1037			China	Commodi	Public Adm	- China - Commodities - Public Administration	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
1038			China	Commodi	Education	- China - Commodities - Education, Health and Other Services	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
1039			China	Commodi	Private Ho	- China - Commodities - Private Households	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
1040			China	Commodi	Others	- China - Commodities - Others	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
1041			China	Commodi	Re-export	- China - Commodities - Re-export & Re-import	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	

Illustration 7 - Example of a By Rank output. This example the same input as the two previous. Selection displays the CO2 emissions caused by the input in China, for different ranks of providers. The first column provides direct emissions (rank 0). Next three columns display ranks from one to three. Fifth column display further ranks, and last column present the total. Highlighted cell highlights that first rank providers (input sectors being rank 0) cause 43.8 kt CO2 emissions in the Chinese Petroleum, Chemicals and Non-metallic minerals sector.

- Eventually, the mode *detailed* uses Equation (VIII) for each chosen indicator. It details therefore the contribution of each item of the final demand on each intermediate

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R
1	***	***	***	***	**	**												
2	***	***	***	***	**	**												
3	***	***	***	***	China	Bangladesh	China	Bangladesh										
4	***	***	***	***	CHN	BGD	CHN	BGD										
5	***	***	***	***	Commodities	Industries	Commodities	Industries										
6	***	***	***	***	Wearing apparel	Textiles and Wearing Apparel	Wearing apparel	Textiles and Wearing Apparel										
7	***	***	***	***	Economic flow (kUS\$ 2013)	Economic flow (kUS\$ 2013)	CO2 emissions (Gg) from EDGAR - Total	CO2 emissions (Gg) from EDGAR - Total										
8	Country	Country	Entity	Sector	** - ** - China - CHN - Com	** - ** - Bangladesh - BGD -	- - China - CHN - Commodities - Wear	- - Bangladesh - BGD - Industries - Te										
9	China	CHN	Commodi	Electricity and steam production	25754.79612	333.6870739	147.8195087	1.915195099										
10	China	CHN	Commodi	Wearing apparel	319377.643	75.19421211	51.83278337	0.012203501										
11	China	CHN	Commodi	Cotton textiles	142421.0074	607.147588	23.67101457	0.100910671										
12	China	CHN	Commodi	Highway freight and passengers	8471.909454	80.9117387	8.240091516	0.07869774										
13	China	CHN	Commodi	Hemp textiles	45141.53066	70.37206343	7.500822246	0.011693187										
14	China	CHN	Commodi	Chemical fibers	29023.53267	93.12477784	7.457358133	0.023927646										
15	China	CHN	Commodi	Petroleum refining	12723.02699	154.1246462	6.58244295	0.079738626										
16	China	CHN	Commodi	Woolen textiles	35535.48336	104.947812	5.949722871	0.017571462										
17	China	CHN	Commodi	Water freight and passengers tra	7058.956072	81.15262787	5.451673321	0.062674652										
18	China	CHN	Commodi	Cement and cement asbestos pr	1408.266216	36.43321685	4.664318856	0.120670466										
19	China	CHN	Commodi	Knitted mills	30549.33872	123.127109	4.306313046	0.017356313										
20	China	CHN	Commodi	Chemicals for special usages	11663.57745	106.101618	3.198035849	0.029091999										
21	China	CHN	Commodi	Raw chemical materials	9781.655736	394.330246	2.429259003	0.097931304										
22	China	CHN	Commodi	Air passenger transport	1114.23742	10.4108274	2.318573541	0.021663488										
23	China	CHN	Commodi	Chemicals for painting, dying and	7619.121897	70.93215266	2.112786818	0.019669526										

input, in indicator's units. It has a row for each sector of the world economy, and a column by item from the final demand and for each indicator. The size of the output could therefore become problematic when the number of items increases, and its use is not advised for wide inputs or numerous indicators.

Note that the *direct impacts* treatment automatically uses a *consumer view* display mode. The choice of one display mode over another relies on a compromise between precision and simplicity. The *detailed mode* allows an in-depth understanding of the interaction between items of the final demand and local externalities. However, it is inopportune for large inputs, and may be complex to analyze. The mode *by rank* offers another interesting insight on a situation, but also provides large and complicated outputs. *Producer* and *Consumers* modes give narrower snapshots of the impacts of some supply chain. The first focuses on the economic sectors which actually produces the externalities, without distinguishing their causes in the final demand array. Conversely, the second allows identifying the direct hotspots of first rank consumption items. Those are useful for having concise information, but do not provide as much information as the formers.

5) Documentation

The flexibility offered by the tool has an obvious cost in its ease of use and of understanding. Most users of the tool are not experts in IO modelling, so its use requires a short introduction to its principle to ensure a proper use. Similarly, the tool could require further updating and maintenance. In that perspective, the deliverable comes with several guides and manuals, to guarantee a sustainable use for PwC's team. First, a report provides a full explanation on the functioning of the tool. The first part of the tool focuses on practical aspects on how and why to use the tool. The second part provides the full report of the construction of the tool and extended discussions on its precision and reliability, close to the content of this chapter. Secondly, a Power Point document offers a concise executive guide to the use of the tool, describing the practical steps for using the tool. Note that the Excel Input file contains a similar operational guide. Eventually, another Power Point document describes step by step the basic maintenance operations.

6) Study case

This study case gives an overview of how the tool may provide relevant information of the footprint of a company. This example will allow us discussing the precision of the tool, and the differences between Full EORA and EORA 26. Let's consider a French restauration brand, which wants to assess its footprint in terms CO₂ emissions. We first select relevant indicators. We have here two choices: CO₂ from Edgar or CO₂ from PRIMAP. Even if PRIMAP is not available in Full EORA, we keep both indicators for further comparison of the results.

The second step is to choose the input mode. We do not look into getting global information about one sector or one country's footprint, but on a precise company. We therefore use the manual input mode, for both models. Third step is defining the input. In a first approximation, we could have a look on the average footprint of the French restauration sector of the size of our company. We therefore choose *Hotels and Restaurant* as the input sector, because it is a close match to our client's activity. For this case, there is unfortunately no more precise sector in Full EORA. Concerning the volume of transaction, we use the annual revenue of the company, which is 40 M€/year before taxes and without subsidies. Eventually, we need to define the display mode. Since it is a first approach, we look forward having some precise information. *Emission view* would be a poor choice, because we would only have a numerical estimation of the total footprint of the company, with no other detail. We have only one sector as input. It is thus equivalent to use the *detailed* or the *production view* mode. However, the most promising display mode in this situation is the *by rank* mode. It will allow us to have

II.A. Overview

a first overview on the propagation of the externalities along the supply chain. We therefore use the *by rank* mode for EORA 26, and the *production view* mode for Full EORA.

	A	B	E	F	G	H	I	J	L
1	Departement	Division	Country Selection	Country match		Sector Selection	Sector Match		Transaction Volume (
2			France	France		Hotel and restaurant services	Hotel and restaurant services		40000
3				#N/A			#REF!		

Illustration 10 - Screen capture of the Full EORA Input tab in the first case. We have a similar input for EORA 26.

First, let's have a look on the numerical values. Edgar provides a footprint of 6.0 ktCO₂ with Full EORA, and 5.6 ktCO₂ for EORA 26. PRIMAP estimates 5.1 ktCO₂. Those estimations are consistent from one to another, with around a 15% difference between the smallest and the biggest estimation. Emission's repartition is also similar: both Full EORA and EORA 26 allocate around 55% of the CO₂ emissions to France, China being the second contributor with 6% and Russia the third with 4-5%. PRIMAP gives slightly different results, with only 48% of emissions happening in France.

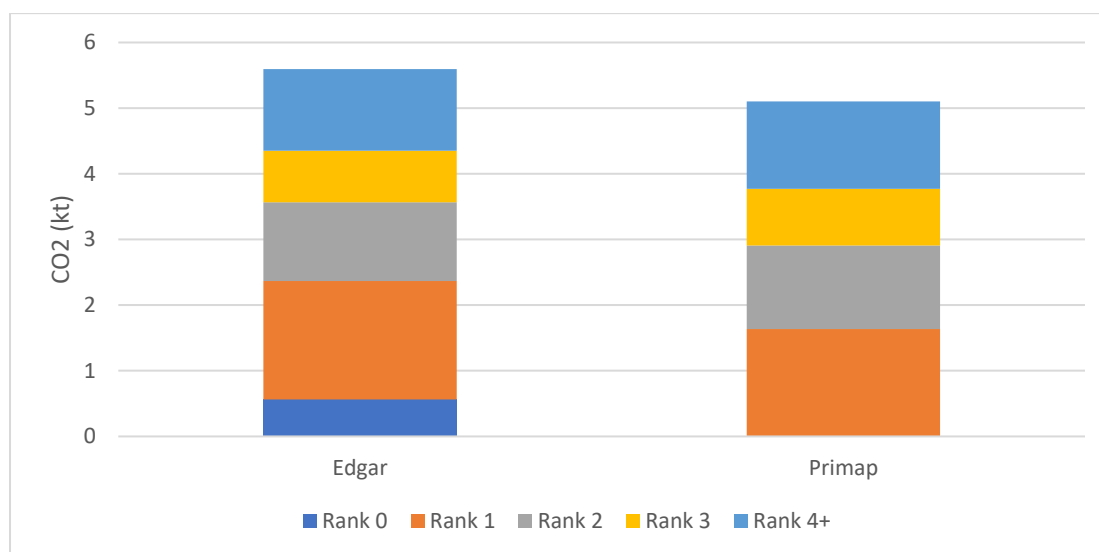


Figure 8- Allocation of CO₂ emissions by rank, EORA 26

The rank repartition provided in Figure 8 highlights that with this first approach, direct emissions are only a fraction of the footprint of the company. Most emissions happen in the first rank, mainly in the “Transport”, “Electricity, gas and water” and “Food & beverages” sectors. In total, both models also place the Chinese “Electricity, Gas and Water” as a major source of indirect CO₂ emissions. Interestingly, this precise source of emissions concentrates largely in high ranks. Both databases count 90% of this sector's emissions in rank 3 or more. The footprint of the brand therefore includes externalities on the Chinese power sector, but only at high ranks in the supply chain. Those do consequently not proceed from direct use of Chinese energy, and neither from Chinese good supplied by Chinese energy, but at least with a third intermediate process. Note that having such insight with a manual approach would probably have required days of work, while it is almost immediate using the tool.

This first estimation already gives us some interesting insights on the externalities caused by our client. However, those results correspond to the averaged impact of the *Hotels and Restaurants* sector. For obvious reasons, our company does probably not have the same characteristics as a so-called average *Hotels and Restaurants* business. We could reach a better level of precision by exploiting more information on the company's supply chain, for instance by using its supply from direct providers. For instance, we consider that the company expenses are 12M€/year in French processed food, 8

M€/year in fresh products from French agriculture, 5M€/year in delivery (by French trucks), and 5 M€/year in Italian processed food.

Country	Sector Full EORA	Sector EORA 26	Transaction volume (k€)
France	Food Products and Beverages	Food & Beverages	12 000
France	Products of agriculture, hunting and related services	Agriculture	8 000
France	Land Transport	Transport	5 000
Italy	Food Products and Beverages	Food & Beverages	5 000

Table 3 - Allocation of providers' information in input table

Table 2 presents the transcription of those expenses items into EORA sector. For three of them, Full EORA and EORA 26 have similar categories. For truck transportation, Full EORA has a more precise sector ("*Land Transport*") than EORA 26 ("*Transport*"). We conserve the same settings for EORA 26, using the *By Rank* display. For Full EORA, we now use the *Detailed Footprint* mode, to discriminate the different input sectors in the results.

Numerical results are now quite different from our first approximation. Full EORA estimates a footprint of 13.5 ktCO₂, while EORA 26 indicates 10.2 kt using Edgar, and 11.4 kt using PRIMAP, which gives us a 10% difference between both environmental sources in Eora 26. Full EORA results scores sensibly higher. It may be partly because of the higher precision for the transport accounting: it causes 60% of the footprint following Full EORA.

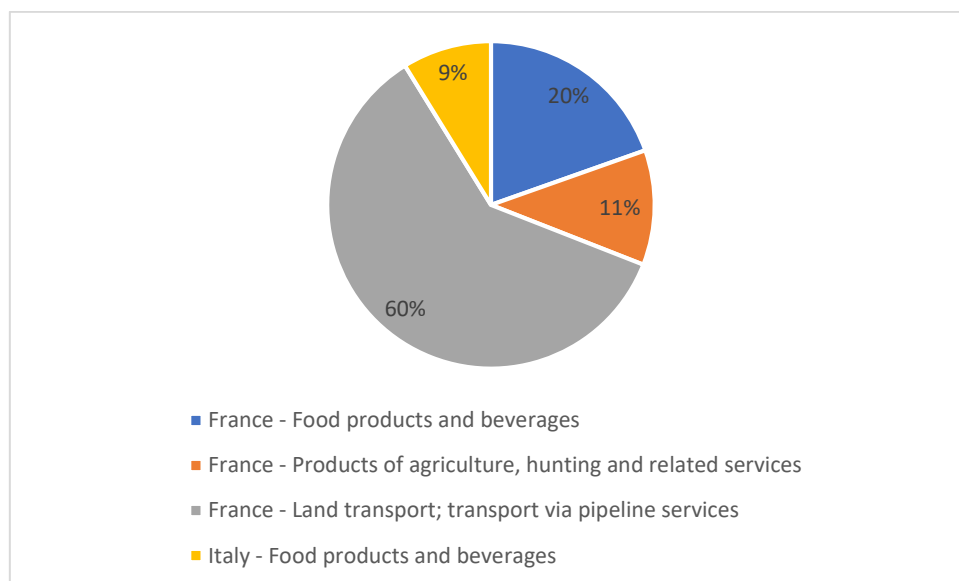


Figure 9 - Repartition of CO₂ emissions among input sectors according to Full Eora

The transport sector is all the more important in the footprint of the company that it is a major direct emitter: both Edgar and PRIMAP acknowledge that the "*Transport*" sector emits most of the rank zero (or scope one)¹ emissions (respectively 82% and 70%).

¹ Those emissions are technically reported as rank zero in the output. However, our input concerns the providers of the company. Direct emissions (or scope zero) of the direct providers therefore correspond to first rank emissions for the company.

II.A. Overview

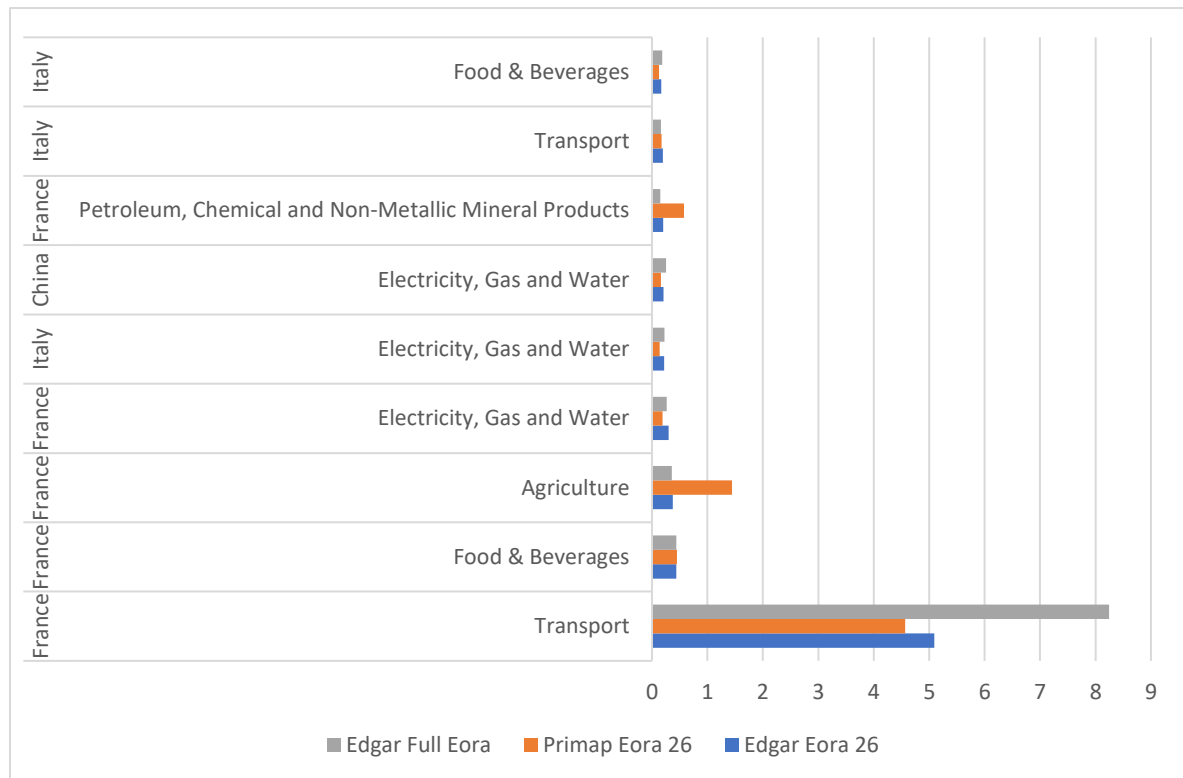


Figure 10 - Repartition of total emissions by sector

Finally, all three estimations agree on the trend of the global repartition of CO₂ emissions among sectors: the French transport sector is by far the first contributor to the footprint of the company. The different sector definition between Full EORA and EORA 26 may explain the large difference in total output. Thanks to more precision in sectors definition, Full EORA provides more accurate results than EORA 26. PRIMAP shows a slightly different pattern than the two Edgar-based estimates, with a lower accounting for Transport emissions at the expense of the Agriculture sector. This difference probably comes from the constitution of the database, as underlines Section I.B.2.

This short study case illustrates a first broad application of the tool to the global assessment of a company's footprint. While the tool enables a quick and easy approach, it also suffers several limits. Previous analysis highlights major characteristics to notice:

- EORA displays broad categories. The footprint of a precise company may differ largely from the average of its economic sector. It is therefore better to use the most precise information possible, for instance the list of direct providers.
- The use of more precise sectors may considerably improve the results. When possible, prefer Full EORA, which offers more sectorial details than EORA 26.
- Both different models (EORA 26 or Full EORA) and different indicators (PRIMAP or Edgar) give different results. Those differences issue from technical and methodological concerns, and apart from the previous remarks, there is no reason to favor one against another (See II.B.3.a). Consequently, one should interpret the results provided by the tool as trends and order of magnitudes better than perfectly exact measures.

II.B. Data, reliability and comparability

The relevance of IO analysis techniques relies fundamentally on the consistence of the associated database. This section analyses different aspects related to the precision of our model. The first- and second-part focus on the characteristics of EORA, respectively for economic and environmental data. The third part tackles the issue of data consistency and variability.

1) Economic data

The tool uses EORA as source database. This database is superior to its competitor on several aspects. First, it is the only available table to cover every country, thus avoiding the use of any *Rest of the World* category. To do so, EORA uses multiple data sources mapping the world economy, using state-the-art numerical techniques to reconcile them and to extrapolate missing data (Lenzen et al. 2012). Secondly, EORA provides the most detailed disaggregation of the world economy, displaying for each country the most precise decomposition available in Full Eora. The counterpart of this feature is the heterogeneity of sectors between countries. The reduced version EORA 26 circumvents this problem by using a common 26 sectors framework. While being less precise, this second model gives more flexibility for further analysis. Academics have free access to the whole database, and widely use it for advanced research (Lenzen et al. 2012; 2013; Alsamawi et al. 2014; T. O. Wiedmann et al. 2015; T. O. Wiedmann, Schandl, and Moran 2015; Brizga, Feng, and Hubacek 2017), supporting regular updating and regular additions of new indicators. Eventually, EORA integrates relatively recent tables of the world economy, on yearly bases.

As introduced in section I.A.7, building a MRIO table requires complex steps of data processing which distort particular values. EORA provides quality assessments of the composition of the table¹, synthesized thereafter. Main results are that while most data are globally consistent with sources, despite low deviations in many cases. However, there is still substantial imbalances and deviations for some sectors, especially for small flows and little countries. EORA also provides comparison with other input output databases², country by country. Different databases provide substantially diverse results for most countries, with differences up to 20% even for big economies. Still, all models appear relatively consistent in terms of order of magnitude, and EORA appears in the normal range of variability. EORA 26 shows however slightly more uncertain results. Eventually, there is no obvious reason to give more credit to a database over another, because their differences proceed from different compromises in their constitution.

Technically speaking, EORA provides time series of the world economy from 1990 and 2015. The full model distinguishes five price layers: basic price, trade margin, transport margin, taxes and subsidies. The basic layer corresponds to the producer price, or the earning perceived by the producer for sealing its products. Consumer price is the sum of all five layers. The use of a MRIO table for environmental analysis requires the use of basic prices (see §I.A.7, *Price Composition*). In consequence, input volumes should exclude all transport and trade margin from intermediates, and taxes and subsidies on trade³. All prices are in thousands of current US dollars. For many countries, Full EORA uses Supply Use tables (SUT), for which there is a distinction between commodities and industries. In this framework, each industry can produce one or multiple commodities. This design has some

¹ <https://worldmr.io/quality/>

² <https://worldmr.io/comparison/>

³ Note that production taxes and subsidies on production account in value-added in the basic price layer.

technical implications but does not change the theoretical base of the analysis. However, this feature impedes the use of the *by rank* display mode in Full EORA¹.

2) Environmental data

EORA embodies a broad satellite matrix, which associates 2729 indicators to each sector of the world economy. Data of this table come from different sources, and do not proceed from the same method than economic data. Most information comes from punctual academic researches based on EORA, for which researchers mapped new data. Updating and documentation is therefore scarcer for this part of the database. Consequently, we proceeded to a detailed examination of all potentially relevant indicators, to assess their consistence.

The matrix includes values for primary energy consumption, GHG emissions, air pollutants, physical extractions, employment, surface, production and yields of agricultural production, nitrogen and phosphorus leakage. However, data quality is heterogeneous between indicators. We ultimately selected fourteen indicators or set of indicators, summarized in the following table. Annex II provides the detailed consistency analysis of those indicators.

Indicator	Description	Source	Unit	Precision
Raw Material Inputs	Raw Material Inputs of biomass, ores, fossil fuels and construction materials	UNEP	Tonnes	Low
Detailed Biomass inputs	Raw biomass inputs in 12 type of plants	UNEP	Tonnes	Medium low
Detailed Ores Inputs	Raw ores inputs by 12 metal categories	UNEP	Tonnes	Low
Detailed Construction Inputs	Raw construction material inputs by 6 categories	UNEP	Tonnes	Low
Detailed Fossil Fuels Inputs	Raw Fossil Fuels inputs by 5 categories	UNEP	Tonnes	Low
Water Requirement	Consumption of water, sorted by: - Blue : Fresh surface and groundwater - Green : Precipitation - Grey water: freshwater required to dilute pollutants	WFN	Mm ³ /year	High in Full EORA Very low in EORA 26
Nitrogen emissions	Emissions of N ₂ O, NH ₃ , NO _x , Exportation to water	(Oita et al. 2016)	Kt	Low
GHG IPCC equ	Global GHG emissions following IPCC 2006 classification	PRIMAP	ktCO ₂ e	Not available in Full EORA High in EORA 26
All GHGs	Emission data for 9 GHG baskets (CO ₂ , N ₂ O, CH ₄ , HFCs, PFCs, SF ₆ , NF ₃ , KyotoGHG)	PRIMAP	kt ktCO ₂ e	Not available in Full EORA High in EORA 26
CO₂	CO ₂ emissions	EDGAR	Kt	High
Primary Energy	Primary Energy Consumption	IEA	Mtoe	Medium

¹ SUTs create additional steps in the interaction between two sectors: instead of having a direct flow from A to B, there is a first flow from sector A to commodity B, and a second one from commodity B to sector C. While this operation required two steps, it only represents one supplementary rank of provider. Since Full EORA mixes both data formats, computing externalities by rank leads to inconsistencies: effect propagation goes by semi steps in countries represented by SUTs, and by normal steps in others.

Primary Energy Mix	Repartition of primary energy sources	IEA	MJ	Medium low
Employment	Number of male and female employees	ILO	'000 full time equivalent	Medium high Not available for most developing countries
Value-added Creation	Sorted by wages, consumption of fixed capital, mixed income, net operating surplus, subsidies and taxes on production	Diverse ¹	'000 Current US\$	Medium

Table 4 - Synthesis of relevant indicators. Appendix II provides the details of the assessment process.

3) Reliability

MRIOT do definitively not offer a perfect representation of the world economy. As illustrated by the integration of PRIMAP data into EORA (See I.B), mapping supplementary data comes also with a lot of assumptions and uncertainties on both raw values and sectorial allocation. In the end, construction of a whole data table triggers many errors and approximation, and one could wonder the credit to give to the results knowing those weaknesses. Quoting George Box: “*All models are wrong, but some are useful*”². This section provides a few elements to circle this issue, by analyzing the variability of the results in between MRIO models.

EEIO's accuracy relies on the precision of its two main constituents: the consistence of the modelling of the world economy, and the association of robust indicators, accurately allocated between sectors and countries. As introduced in paragraph I.B.1.d, several MRIO databases exist, with different choices in source prioritization, sectorial disaggregation and data processing. The emergence of those competitive databases fuels a growing literature of the analysis of divergence between model – and subsequently of their global accuracy. Next paragraph introduces the main insights of this burning research question. However, note that this issue is still an open question in the scientific community. Improvements in data sources, numerical process and methodology may further reduce those uncertainties.

Both economic modelling and environmental externalities allocation create divergences in model comparison. (Moran and Wood 2014) highlight that both dimensions create a comparable distortion in terms of results. Their methodology relies on comparing Consumption Based Accounting between five models at country level, with and without environmental data harmonization. Typical difference lies in a 5-30% range between all models, after harmonization of environmental data³. Nonetheless, models capture similar trend evolution over years. MRIO approaches therefore appear relatively consistent from one to another, while showing major divergences in absolute values. (Owen et al. 2014) arrives to a similar conclusion, underlining the influence of both final demand, technical coefficient and environmental data differences. The issue on the relative credibility to give to one model over another is tricky, because they issue from divergences in methodological approach. Nevertheless, (Steen-Olsen et al. 2014) supports that higher sectorial disaggregation always deliver superior results for carbon accounting applications. The main argument is that data providers aggregate economic sectors not on environmental characteristics (such as carbon intensity), but on economic similarities. Consequently, broad categories gather industries with very different characteristics, averaging the environmental impact over subsectors. That is the case for example with

¹ Value-added indicators belong to the economic core of EORA. Their computation comes from the same process than the rest of the table. The use of those indicators therefore rely on the same issues flagged in section II.B.1.

² (Box and Draper 1987)

³ See again <https://worldmrio.com/comparison/> for a country by country comparison.

livestock in agricultural sectors, aluminum processing in metal industry or cement production. Same observation also apply for raw material accounting (de Koning et al. 2015). This critic supports the creation of footprint-aimed MRIO, where sectorial aggregation reflects environmental similarities. EXIOBASE follows this philosophy, by focusing on technological modelling (Wood et al. 2014). However, this process is very time consuming. Therefore, such database covers few countries, and relies on sensibly older data.

This quick literature review brings interesting insights for the use of our tool. First, it is obvious that there is no silver bullet in terms of consumption-based accounting. MRIO is a top down modelling approach, which require sensible simplification (therefore distortion) of reality. Results provided by the tool do not perfectly capture the exact externalities of an activity, but better offer an outline of the order of magnitude and the repartition of those externalities. While MRIO models differ widely in quantitative results, they still agree on global trends, both through time and between countries. There is therefore a strong indication on the global relevance of the method. The tool relies on both EORA 26 and Full EORA. Literature highlights that both model display sensibly different results (Moran and Wood 2014). When comparing those results, user should keep in mind that while being related in their conception; those two models propose a different picture of the world economy. Results are therefore comparable in the range of the intermodal variability previously introduced and could sometimes slightly disagree. While Full EORA provides in theory better results than EORA 26 (by construction and thanks to higher disaggregation), the tool offers the use of both because both could be useful, depending on the aim of the study. Some indicators are available in EORA 26 and not in Full EORA, and reciprocally. Nonetheless, some new evidences tend to reduce the gap between EORA 26 and Full EORA. Divergences in country-by-country analyses seem to reduce in most cases when harmonizing environmental data. Similarly, newly allocated PRIMAP in EORA 26 presents very close results to Edgar as informed in Full EORA. Therefore, it may be that differences between both models proceed in a large way from poor environmental data allocation in EORA 26. Table 4 therefore includes a concise appreciation of the quality of the environmental mapping for each model, in order to enlightens the choice of one model over the other for every choice.

4) Compatibility with Life Cycle Analysis

Life-Cycle Analysis (LCA) is the main approach for assessing the externalities associated with the supply chain, the use and the disposal of a product. In many ways, EEIO appears comparable to LCA: both aim assessing the externalities embodied with previous processes of supply chains. Because of practical and theoretical differences, EEIO and LCA are however not exactly equivalent, and may provide sensibly different results. This section discusses the comparative characteristics of both approaches, highlighting their respective forces and drawbacks. The first paragraph introduces the reader to the basics of LCA. The second one provides a comparative analysis of both and offers good practice guidelines. The last part sketches a set of techniques to reconcile both approaches.

a) *A short introduction to LCA*

LCA designates a set of techniques to measure all environmental effects caused by a product, at every step of its life cycle. Norms ISO 14040 and 14044 defines the standards for such assessment. The three main steps of a LCA are:

- Defining the goal and the scope of the study
- Making the inventory of all flows at every step of the life cycle (LCI)
- Assessing the environmental impacts of those flows

The Life Cycle Inventory (LCI) aims at building a representation of all processes involved in the life cycle of a country, from the extraction of raw material to the packaging of the final product and

potentially including its use and disposal. The standard approach is the Process-Based Analysis (PBA). The idea is to manually reconstruct all the individual processes involved in the life cycle of a product. This approach relies on both process databases and contextual surveys, which provide the flows in and out specific industrial processes. The main advantage of this approach lies in its specificity. User could precisely characterize the industrial processes, and therefore highlight the consequences of specific choices over some others. The main drawbacks of PBA come from the manual aspect: user must reconstruct process by process the life cycle of a product, which is very time consuming. Similarly, such reconstruction cannot be exhaustive: user must draw the perimeter of the process and must exclude flows when they become supposedly too small to be significant. Such criterion applies most of the time to the absolute size of the flows or to the rank in the supply chain. This so-called *truncation error* systematically hinders a portion of the externalities of the LCA. Eventually, PBA is also subject to data reliability concerns: different databases offers different estimations of the externalities of industrial processes. This limits the repeatability of such analysis.

b) Comparative analysis

While PBA offers a narrower scope to tackle the issue of externalities assessment, EEIO could provide comparable insights on such problem. A standard approach is using the suppliers of the different parts of the product under study to fill the final demand array Y^1 . While PBA is a *bottom-up* approach which builds the global picture from individual industrial processes, EEIO uses a *top-down* reasoning, assuming the specificity of the supply chain from macro-economic considerations. The two techniques use therefore opposite approaches. The results provided by both techniques may disagree quite strongly. Early tests of the tool at PwC on actual cases highlighted quantitative difference ranging up to 50%. The analysis of this disagreement is a tricky problem, because it may root in many factors: data variability, methodology, characteristics of the study case... However, the next paragraph explores the main cause of the disagreement between PBA and EEIO.

Apart from its practical limits, EEIO offers a consistent framework for analyzing the global externalities of a product, at least in theory. Indeed, an EEIO table containing all industrial processes and perfect data provides a consistent and exhaustive approach to the externalities of an activity. The construction of such ideal EEIO tables therefore allows assessing the pure correctness of the PBA methodology. Studies using this kind of approach highlight that the *truncation error* previously introduced could hinders in theory as much as 50% of the global externalities of a product (Lenzen 2000; Kennelly, Berners-Lee, and Hewitt 2019). In parallel, some authors underline that the approximations related to sectorial aggregation may overweight the *truncation error* (Yang, Heijungs, and Brandão 2017), while some other argue that such situation is unrealistic (Pomponi and Lenzen 2018). In the end, a global overview of this controversy underlines that the *truncation error* always exists, is only weakly dependent to the settings of the PBA and may not be precisely quantified (Ward et al. 2018). In other words, the PBA approach may give more personalized results, but it always underestimates by a significant but unknown margin the real externalities caused by a product. On the other hand, EEIO is a macro-economic approach, that inaccurately captures specific processes.

In the end, both approaches rely on contestable modelling of the functioning of real-world supply chains. There are therefore no *right* or *wrong* results, but two approximation of an unknown and uncachable reality. The only obvious result is that PBA always underestimates the externalities due to the truncation error. The operational consequence of this assessment is that PBA and EEIO are

¹ This approach is consistent with a *cradle-to-gate* analysis, which excludes the analysis of the externalities associated with the use and the end of life of the product. There are however some other techniques to include those further steps in the analysis. The idea is to account for those supplementary steps by the mean of their indirect requirements. For instance, the electricity or fuel consumed while using the product or for its disposal.

not equivalent and must not be compared side to side. Practically, the choice of one approach over the other depends on a compromise between different parameters, summarized in table 5.

	Process Based Analysis	Input Output Approach
Precision	Process-specific	Aggregation on broad economic sectors
Time Requirement	Requires extensive manual work	Low, relies on automated processes
Exhaustiveness	Low: subject to truncation error	Complete in theory
Input data requirement	Requires precise data on material inputs	Requires only monetary flows
Sensibility	May capture specific situations	Captures only global figures

Table 5 - Comparative analysis of PBA and EEIO applied to LCA

c) Toward Hybrid Life Cycle Analysis

PBA and EEIO uses two opposed approach to assess the externalities associated with the life cycle of a product. Both approaches have forces and drawbacks, and it may be unclear to favor one against the other. However, early 21st century saw the emergence of a class of hybrid techniques, aiming to reconcile the completeness and efficiency of input output approach with the technological detail and precision of Process Based Accounting. The content of the following paragraph summarizes the content of two recent literature reviews on this topic (Islam, Ponnambalam, and Lam 2016; Crawford et al. 2018).

A sound and easy way to combine PBA and EEIO analysis is the *tiered* approach. It consists in applying a standard LCA approach, while filling the unexplored flows with EEIO estimations. This method allows estimating portions of the supply chain that previously constituted the *truncation error*. It offers great flexibility: user could decide which process step are worth proceeding to a detailed process analysis and for which an EEIO analysis is enough. Such criterion could be horizontal or vertical. In the first case the user chooses to assess via EEIO all flows inferior to a certain threshold, while the second choice consists in choosing the method to apply depending on the rank in the supply chain.

Conversely, one could first apply a global IO analysis to identify the main sectors embodying externalities and explore then the most significant branches with PBA: this is the *path-exchange* method. Such operation is repeatable: user could specify the previously identified process with a traditional PBA, and then computing new estimations for the following rank using the EEIO approach, exploring step by step the main branches of the supply chain, refining at each round the previous estimation. Definition of the boundary between one approach and the other permit a flexible choice of the required level of precision, without threatening the global consistency of the assessment.

Considering drawbacks, both approaches may meet three limitations related to interoperability between EEIO and PBA. First, the EEIO tool requires monetary values as inputs, while PBA uses material flows. User should therefore ensure the proper equivalence between both. Similarly, differences in sectorial disaggregation may limit the efficiency of hybrid approach, by impeding IO analysis from exploiting all information provided by PB LCA (Kennelly, Berners-Lee, and Hewitt 2019). Eventually, the tool includes a limited number of indicators. The application of hybrid techniques could then only apply to common indicators for both methods. Some assumptions could however tackle this last issue. One could compute for each common indicator the rate of externalities volume covered by the PBA method compared to externalities estimated by EEIO method, and apply this rate to extrapolate the indicators non available in the tool.

There are other types of hybrid techniques, but they require modification of the IO table, and do not cope well with usual PBA databases. Therefore, they do not attract much attention outside of the academic community. *Matrix Augmentation* aims increasing the sectorial disaggregation of the table using process specific data. For instance, one could add a biofuel sector one the petroleum sector, using different inputs than the basic sector (Malik et al. 2015). The global assessment finally exploits this refined table as a conventional MRIO table. Another method is *Integrated Hybrid*. It relies on the creation of a process matrix, which combines mathematically with the IO table. Both methods are substantially more complicated than the *tiered* and *path-exchange* approaches, and are not easily applicable with the tool.

II.C. Technical Implementation

The tool runs on locally maintained files. Transmission of input and output data uses excel files, allowing a flexible exploitation. A single script allows basic maintenance operations. Figure 11 provides an overview of the architecture of the program.

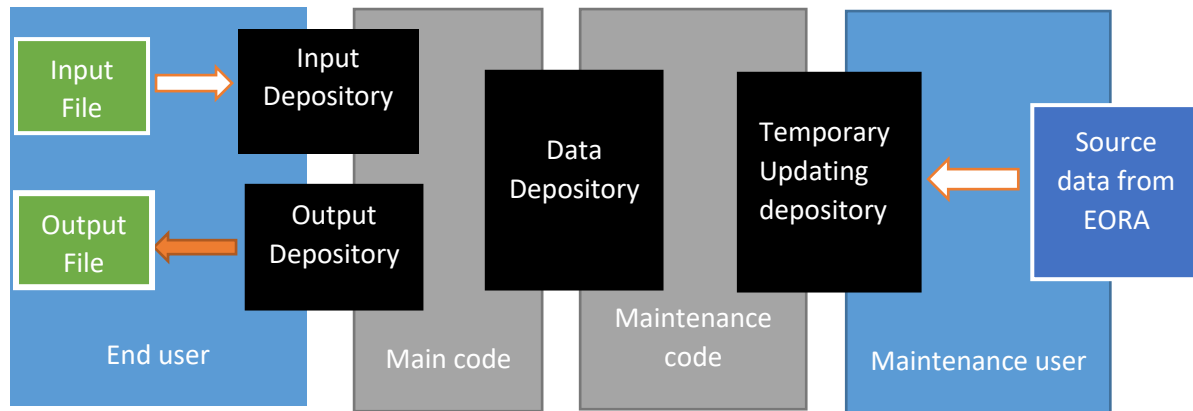


Figure 11 - Functional organization of the tool implementation. Black boxes are depositories on the cloud. Full arrows represent downloading from the cloud. Blank arrows are uploads to the cloud. Light blue boxes advocate for external data sources.

Following paragraphs detail the implementation of the tool. The first section describes the operational algorithms, the second the source data organization. Section three highlights the functioning of the interface. Fourth section focuses on the distribution and accessibility of the program. Eventually, the last paragraph details the maintenance operations.

1) Main codes

There are three scripts supporting the tool, all of them running on Python 3.7. This section focuses on the two main scripts, *Main* and *eora_treatment*¹.

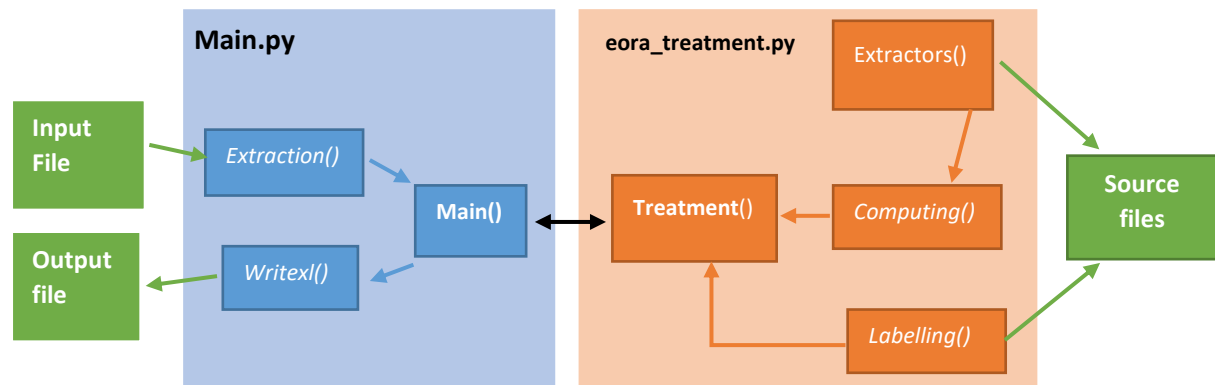


Figure 12 - Functional scheme of the treatment process

Figure 12 illustrates the global treatment process. *Main* manages the interface with the input file, while *eora_treatment* proceeds to actual computation. The main script is the executable, via the main function. It manages the following tasks:

- Extracting and formatting commands and data from the input file
- Organizing and writing output data in the output file
- Calling the *eora_treatment* script

¹ We describe the third script (*maintenance_script*) in subsection 5.

The *eora_treatment* script manages all actual operations on the matrices, and the logical interpretation of the commands. It contains five types of function, corresponding to different tasks:

- *Treatment* interpret command data and manage the execution of relevant functions
- *Input computers* build input arrays for treatment modes with non-manual inputs.
- *Allocations* apply the relevant operations on the matrices.
- *Labellers* manage the labelling of raw outputs
- *Extractors* load EORA matrix and extract relevant lines/columns.

The main script calls the *Treatment* function to interpret input data. It treats successively EORA 26 and Full EORA commands, using an identical process. A succession of logical gates allows applying the good functions for both input data transformation, treatment, and output labelling. Note that while both EORA 26 and Full EORA rely on the execution of the same functions, small differences in data formatting causes some slight differences in the code. To allow this distinction with minimal effort in the code, all functions use a *runid* argument as a marker of the current model under treatment. The arguments values zero when operations concern EORA 26, and one for Full EORA. Successive operations include generating input data for national footprint and industry footprint modes, computing the raw output and adding labels to the output.

Treatment mode/Display mode	<i>Detailed</i>	<i>Consumer view</i>	<i>Production view</i>	<i>By rank</i>
Direct Impact	/	Direct impact <i>Simple labelling</i>	/	/
Manual Input	<i>Detailed footprint</i> <i>Complex labelling</i>	<i>Sectorial footprint</i> <i>Simple labelling</i>	<i>Sectorial Footprint</i> <i>– emissions side</i> <i>Simple labelling*</i>	<i>Footprint by rank</i> <i>By Rank labelling</i>
Industry Footprint	Demand Industry <i>Detailed Footprint</i> <i>Complex Labelling</i>	Demand Industry <i>Sectorial footprint</i> <i>Simple labelling</i>	Demand Industry <i>Sectorial Footprint</i> <i>– emissions side</i> <i>Simple labelling*</i>	Demand Industry <i>Footprint by rank</i> <i>By rank labelling</i>
National Footprint	National Demand <i>Detailed Footprint</i> <i>Complex Labelling</i>	National Demand <i>Sectorial footprint</i> <i>Simple labelling</i>	National Demand <i>Sectorial Footprint</i> <i>– emissions side</i> <i>Simple labelling*</i>	National Demand <i>Footprint by rank</i> <i>By rank labelling</i>
No treatment	/	/	/	/

Table 6 - Treatment operations in function of the selected treatment mode and display mode.

* Consumer view and production view call the same labelling function, but with different arguments. Consumer view only considers sectors in input as arguments for the labelling; while the production view process considers every sector of the world economy.

Table 6 summarizes the treatment applied for each input combination. The treatment mode mainly determines the treatment to apply to input data. Display mode influences both the treatment operations and the labelling process.

When calling *EORA_treatment*, the *main* function shares the following information:

- *Sectors26*: list of two lists, containing in first position the list of EORA 26 sectors (represented by numerical ids), and the associated transaction volume in second position.
- *Sectors_full*: same as previous, but for Full EORA
- *Indicators*: list of two lists. The first one embodies the id of each selected environmental indicator. The second one carries ids of value-added indicators

II.C. Technical Implementation

- *Report*: list of strings. Contains all information that will form the treatment report. Integrates additional information at each step of the process
- *Div26*: the optional list of divisions associated with the sectors for EORA 26. Note that production view and by scope display do not integrate such information
- *Divfull*: same for Full EORA
- *Year*: Year for source data (currently limited to 2015)
- *Country*: EORA id of the country selected for the national footprint mode.

Conversely, *EORA_treatment* returns the following data:

- *Report*: information on the treatment process, transcript of the output in the treatment console
- *Output*: a list of two numpy matrixes, containing the output for EORA 26 in first position, and for full EORA in second position.
- *Data_in*: a list of two numpy matrixes, including information on the actual input considered for the computation (including potential treatment).

The freezing of the Python script uses Pyinstaller¹, on standalone file mode (-F). This software packages the whole code in a unique executable, which does not require an installation of Python to run. Changes in operative systems may however limit its operability.

2) Data repositories

For enhanced transportability, both scripts use relative paths to access source files. Those paths belong the preamble of *Main* script. Concerning *eora_Treatment*, the first part of the paths are in preamble, and the terminations are in the *path_matrix* function. Any modification in the organization of source files would thus require updating the path in the script.

Source path	Data	Code	
		EORA 26	Data
		Full EORA	A
			Index
			Q
			V
			Y
			Leontief_inversion
	Input		
	Output		

Figure 11 - Data deposit organization

Data stored in the deposits require substantial modification from source files provided by EORA. In total, the tool requires the storage of nine files for each model:

- Technical matrix *A*
- Leontief inverse *L*
- Final demand array *Y*
- Gross output array *G*
- Environmental matrix *Q*
- Normalized environmental matrix *q*
- Normalized value-added array *v*

¹ <https://www.pyinstaller.org/>

- Environmental matrix labels
- Trade matrix labels

Those files are quite heavy: each year weights in total slightly less than 15 Go. The maintenance algorithm can automatically compute those files (see section 5).

3) Interface

The interface between the user and the tool relies on a two Excel files. The first one provides takes both the commands and input data, while the second displays the results computed by the algorithm. The first part of the document already introduced the guiding lines for using those files. Further paragraphs therefore focus on the technical designs supporting them.

a) Interpretation of input data

The interpretation of input data uses the *xlsxwriter* library, while the writing function uses *xlrd*. Both libraries picture an excel file as a list of two-dimensional lists, each sub-list being a different tab. The reading and the writing of both files relies on the absolute position of the relevant cells and columns. Addition of supplementary columns in the input file would therefore require the adaptation of the reader's code. For *EORA 26* and *Full EORA* tabs, the algorithm scans all the attributes of a single item before analyzing the next one. The attributes are the id of the sector, the monetary flow and potential organization information. The scanning stops when the reader encounters an unexpected value for the *sector id* (not integer nor string). The construction of this last identifier involves a *vlookup* Excel function. It returns therefore an error value whenever a country id or a sector id misses, ensuring that the scanning terminates at the line at which user stopped adding correct entries. Same principle applies for interpreting the list of selected indicators on the *commands* tab.

The transformation of literal sectors and country name into EORA ids takes place directly on the input file. For *EORA 26*, the transformation is straightforward: columns M and N contains the list of EORA sectors, with their associated order in *EORA 26*. Columns P and Q contain the same info for countries. A *VLookup* function then associates to each selected sector the corresponding id. The computation of the equivalent EORA id eventually associates both ids, where e is the EORA id, c the country id and s the sector id:

$$e = 26 * c + s$$

Both the sector heterogeneity and the use of Supply Use Tables in *Full EORA* makes the association more challenging. We overcome it by adding a hidden tab to the file, named *Index*. This tab proceeds from the source label file, where each couple country/sector corresponds to a unique index. By aggregating the sector and country name, we associate each index with a unique string, describing both the country and the sector. The *Full_EORA* tab builds a similar aggregated name, and a *VLookup* function ultimately associates the input string with the corresponding index. Countries represented by SUT show both industries and commodities. The mapping only considers the commodities, because they correspond to the actual goods and services being traded. A similar system applies for interpreting the indicators and parameters choices.

Eventually, EORA data currency is kUS\$. Most application of the tool will however take input in current €. A rudimentary function converts the last in US\$ for the data year. The principle is to use the current conversion rate to convert 2018 € into 2018 US\$. The second step is subtracting the cumulated inflation, to reflect the monetary flow associated with the same material quantities on the year of the database. However, some cases may require other conversion, be it other currencies or different years. User can therefore parametrize the value of inflation and conversion rate.

b) *Assisted selection*

Most input fields provide rolling lists for value selection. The building of most lists is quite straightforward¹. However, the heterogeneity in Full EORA sector requires a more advanced solution. We therefore use a system of chained rolled list², using a hidden tab (*List*). This tab displays, for each individual country the list of sectors available. When the user chooses a country, the rolling list of the sector selection automatically updates, and displays the sectors available for the chosen country.

The input file provides another assisting process, to allow easier integration of large data inputs. User can indeed past a list of countries/sectors into the corresponding column, instead of manually selecting each country and each sector. Depending on the data source, it may happen that country and sectors from the source do not exactly match EORA's denomination (ex: People's Republic of China instead of China...). A simple *VLookup* function therefore applies to the input data and associates an existing EORA country/sector serves to the entry. However, this association is not always relevant. In order to ease the check of those matchs, the file automatically provides an indicator of the state of the association: green if there is the exact same denomination in EORA, orange if there is a match but not exact and red if there is no match at all. User could then focus on orange and red items only, ensuring that the firsts are relevant, and manually associating the seconds. Technically speaking, the validation system relies on a simple equality test, between the input entry and the associated value, combined with a conditional formatting setting. When both entries are equal, test returns *True*, which automatically colors the box in green. When entries are not equal, test returns *False* and the box turns orange. Eventually, when one entry is missing or when the *VLookup* function does not return anything, the test returns an error, which turns the box red.

c) *Output file*

The output file includes four tabs, displaying respectively the treatment report, the output for EORA 26 and for Full EORA, and the transcription of data used as input for the computations. Those tabs are direct transcription of the raw data from the tool. Concerning the different display modes for the output, their design aims finding a compromise between completeness and clarity. The use of a pivot table allows the user distinguishing easily the repartition of flows among countries or sector. The raw output could also support advanced visualization, using for example *Tableau* or *Power BI*.

4) *Delivery*

The tool holds in a compressed folder of 3.2 GB, which includes all supporting files, the executable app and documentation. The documentation includes a global presentation of the tool, a user guide and a maintenance guide. Considering both the weight of the supporting files and the computational requirements of the executable, a remote access is currently under implementation. The principle is to upload the application on a virtual machine. The utilization of the tool follows the exact same principle as in the local distribution, but the execution takes place on an external server. Apart for enhancing the ease and speed of use of the tool, such solution allows for centralized maintenance and upgrading. The execution of this operation is under the supervision of PwC's teams.

5) *Update routine*

This last subsection explores the technical insights associated with the regular maintenance and improvement of the tool. Maintenance operations could take place whenever the EORA team provides actualized data or new tables. In case there is no change in data organization, an automatized

¹ <https://support.microsoft.com/fr-tn/help/4027045/excel-create-a-drop-down-list-in-excel>

² <https://www.commentcamarche.net/faq/33905-listes-deroulantes-de-cellule-en-cascade> (In French)

procedure manages most operations. Nevertheless, deeper changes in the matrices may require further modifications of the tool.

a) *Automatized process*

The *maintenance_script* manages almost all the required operations required to generate source files using raw data from EORA. It still requires a preliminary manual step consisting in downloading the source file from EORA website¹. User should choose the basic price tables (noted *bp* for EORA 26). After saving the tables, user must check the source path and individual file names in the maintenance script, along with the destination path. Once done, the *update_year* function will automatically create the computed files in the destination depositories. The function requires two arguments: the year to update, and a threshold, in kUS\$. This threshold is an arbitrary level used to clean the trade matrix, by setting to zero each coefficient under the threshold (by default 10k\$). Indeed, EORA balancing processes tend to create improper noise on the matrix that could unduly enhance the complexity of the calculations. However, those inconsistencies have a very small amplitude, and have an imperceptible effect on the final output. The script also erases a bench of inconsistent lines and column in EORA 26 (Sudan, South Sudan, Former Ussr, and Rest of the World). In detail, the script computes all matrices required for the use of the tool (see C.2. Data repositories). When uploading new data, EORA's team often updates older tables, using most recent data. We therefore recommend updating all the tables on regular basis. When doing so, user should also update the *eora_version* variable value in the *main* script, to keep track of changes.

b) *Deeper modifications*

EORA aims providing consistent time series. This implies that data structure should remain constant over time in theory. However, it may happen, that some part of the matrix undergoes substantial changes (especially for the environmental matrix). Such change would imperatively require some changes in the code of the tool. The first step is to track the modifications in the structure, by comparing previous data labels with the new version. If there is a change for instance in the list of environmental indicators, user has to re-index the indicator selection tool of the input file. The *update_year* algorithm should however keep working without any modification. Note that such modification could also allow adding some other indicators to the selection set. One could do so by simply adding new lines on the selection set table with the new indicator and its associated indicator, and by extending both the *VLookUp* function and the rolling list.

Another possible modification could concern sector's list. While this is least probable considering the time series consistencies objective of EORA team, this would require substantial modifications of the code. First, there is a few operations in both *eora_treatment* and *maintenance_script* files that use the size of the trade matrix. User should therefore update this value on each of those scripts. The second set of modification would concern the adaptation of the module translating the written input into an EORA id in the input file (see C.3. Interpretation of input data). Eventually, new label files should replace the former in the data depository, because the *maintenance_script* does not update label files.

A last possible change in raw data concerns the formatting. Current files of EORA 26 are *.txt* files, with '\t' as delimiter; and Full EORA files are *.csv* with ',' as delimiter. Changes in delimiters and files formatting require changing the delimiter arguments in the *np.loadtxt* of the *maintenance_script*. Changes in the label files format would also require changes in the label extractor functions.

¹ <https://worldmrio.com/EORA26/> for EORA 26 and <https://worldmrio.com/EORA/> for Full EORA. The access to Full EORA data requires to connect using the commercial license identifiers.

II.D. Synthesis and perspectives

The tool offers a direct application of the principles of EEIO analysis. It constitutes a quick and efficient way to measure the externalities of various human activities on several social and environmental indicators. The interface allows a flexible use of the tool, paving the way for various application cases. Even if its numerical and sectorial precision is subject to discussion, the tool uses the best techniques available today to provide insights on the global footprint of economic activities. In that perspective, it constitutes a strong asset in PwC's toolbox for analyzing the environmental performance of economic actors, but also for life cycle analysis, risk exposure assessment, or externalities' investments of investment choices. Being entirely developed in the company, the tool could support any further modification and extension.

Practically speaking, the state of the tool at the end of the internship meets the initial target. It is now fully operational, with sensible evidences on its validity and precision. However, two main aspects appear incompletely developed. First, the deployment of the tool is not fully done. The installation of the tool on a virtual machine would definitively ease its use, increases its performance and allows centralized maintenance. This implementation is currently under implementation by the informatic support team of PwC. The second unexplored topic is the issue of data visualization. The tool provides the results as extended tables, which require substantial analysis and retreatment. The development of pre-defined visualizations of the results would arguably increase the attractiveness of the tool.

In a broader perspective, EEIO is a very dynamic topic in the scientific community. This field of research has undergone tremendous progress in the last twenty years and will continue to grow during forthcoming years. There are good chances that the current drawbacks of databases will continuously reduce, while new applications will appear. This is particularly true for EORA, which is favored by many academics all around the world. This interest of the scientific community will probably allow its continuous improvement, especially in terms of available indicators (air pollution, deforestation, land use...). Further evolutions could also embody a merge with LCA tools, to ease the application of hybrid life cycle analysis methods. Another very promising topic concerns social externalities. Eora provides a few socio-economic indicator (employment, wages), which could support the application of social right indicators. Two ways of making appear consistent with such extension. The first is to add specific indicators to the raw matrix. This would require substantial work, because it supposes to seek and assign data to each economic sector of every country of the world. Alternatively, one could combine the available employment indicator with indicators of social risks at national scale. Although more convenient to apply, such approach would hinder the qualitative differences between economic sectors.

Eventually, the development of this tool highlights that EEIO is mature for commercial applications. Public awareness pushes economic actors to master their global externalities. EEIO provides a flexible and efficient approach to tackle this issue. While numerical precision remains imperfect, this class of model appears nowadays sufficiently robust to provide powerful insights to business- and policy-makers. Eventually, EEIO could harmoniously integrate with pre-existing techniques to deliver state-of-art assessments on operational decisions.

CHAPTER III: FORECASTING PLAUSIBLE EVOLUTIONS OF THE WORLD ECONOMY

MRIO tables provide a snapshot about the current interlinkages in the world economy. However, sustainable decision-making does not only require insights on the present state of commercial networks, but also on their possible evolution. Indeed, commercial networks have the property to propagate shocks throughout the world economy (van der Veen 2004; Okuyama and Santos 2014; Otto et al. 2017), including at the aftermath of extreme climate events (Willner, Otto, and Levermann 2018). With climate change increasing both the intensity and the frequency of extreme weather events, the stakes are high for anticipating the exposure of supply chains to climate change (Levermann 2014). Such assessment requires both characterizing future's climate and having quantitative insights on the forthcoming state of interconnexions of the world economy. While the former is already a major topic in the modelling community (IPCC 2012), the last attracted yet much less attention.

This last chapter introduces a promising method to meet this challenge, by forecasting MRIO tables using macro-economic projections. This original research took place at PIK, under the direct supervision of Leonie Wenz¹. This chapter proceeds from the direct outcomes of this research project, which will continue in the following months. The first part extends this introduction and provides some new notions. The second part introduces the method and its implementation. The third part provides a critique evaluation of the outcomes of the model. The fourth concludes on the current state of the project and draw several evolution perspectives.

¹ leonie.wenz@pik-potsdam.de

III.A. Contextualization

Illustrating both technological and trade relations, MRIO is a privileged tool to study interlinkages in commercial networks and the propagation of shocks throughout supply chains (Otto et al. 2017). They rely on nationwide economic surveys, which requires a few years of delay for data gathering and balancing. Nowadays several techniques exist to bridge the gap between past tables and current year, using whether constrained optimization or biproportionnal balancing (Lahr and de Mesnard 2004; Lenzen, Gallego, and Wood 2009; Miller and Blair 2009). Both approaches rely on mathematical processes to minimize an underdetermined problem, under some metric. However, those techniques focus on capturing recent trends, and does not include prospective simulation. At our knowledge, the few attempts to forecast MRIO tables rely on integrated macroeconomic models, such as I3ME¹ (Mercure et al. 2018) and GINFORS² (Distelkamp and Meyer 2019). Among the drawbacks of such integrated approach, results depend highly on theoretical approach of the world economy (Meyer and Ahlert 2019) and on parametrization choices. Moreover, those techniques work in a closed environment and integrate poorly in broader frameworks. Our method aims at bridging this gap, by providing middle term projections of the world economy using a flexible and open approach, consistent with different uses.

The Shared Socioeconomic Pathways (SSPs) provide a framework to build such projections on. This concept aims to ground a body of literature to document the societal and ecosystemic consequences resulting from five alternative social, economic and environmental development path, in the absence of climate change and climate policies (O'Neill et al. 2017). Initially, qualitative assumptions outline the five narrative in various dimensions: Demographics, Human Development, Economy and Lifestyle, Policies and Institutions, Technology and Environment and Natural Ressources. The International Institute for Applied System Analysis (IIASA) gathers numerical interpretation of SSPs scenarios as a common framework for the modelling community (Riahi et al. 2017) and currently includes population (Kc and Lutz 2017), and GDP projections (Dellink et al. 2017; Leimbach et al. 2017; Crespo Cuaresma 2017). Our work aims to engage in this framework, by providing quantitative insights of the global trade network under different narratives.

We therefore define an iterative method to update MRIO tables following macroeconomic drivers for the world economy. First, we project the near-term evolution the final demand, using GDP growth rate as a driver, and we use the algebraic properties of the trade matrix to compute the output that each sector should produce to meet the projected macroeconomic conditions. Secondly, we apply a biproportionale technique (RAS) to adjust trade flows to the new state of the world economy. The process repeats then using the newly obtained table as a basis for the next time step. Our approach thus combines well-established IO techniques in a comprehensive sequence to capture both qualitative and quantitative evolutions of global trade, in line with an exogenous macroeconomic pathway.

¹ <https://www.e3me.com/>

² <https://www.gws-os.com/de/index.php/global-developments-and-resources/models.html>

III.2. Method: an algorithm to project future trade relations

III.2. Method: an algorithm to project future trade relations

Our algorithm produces a time series of MRIO tables, using a known table as base year and iterating a two-step process over the years. Figure 13 illustrates the basic principle of the method.

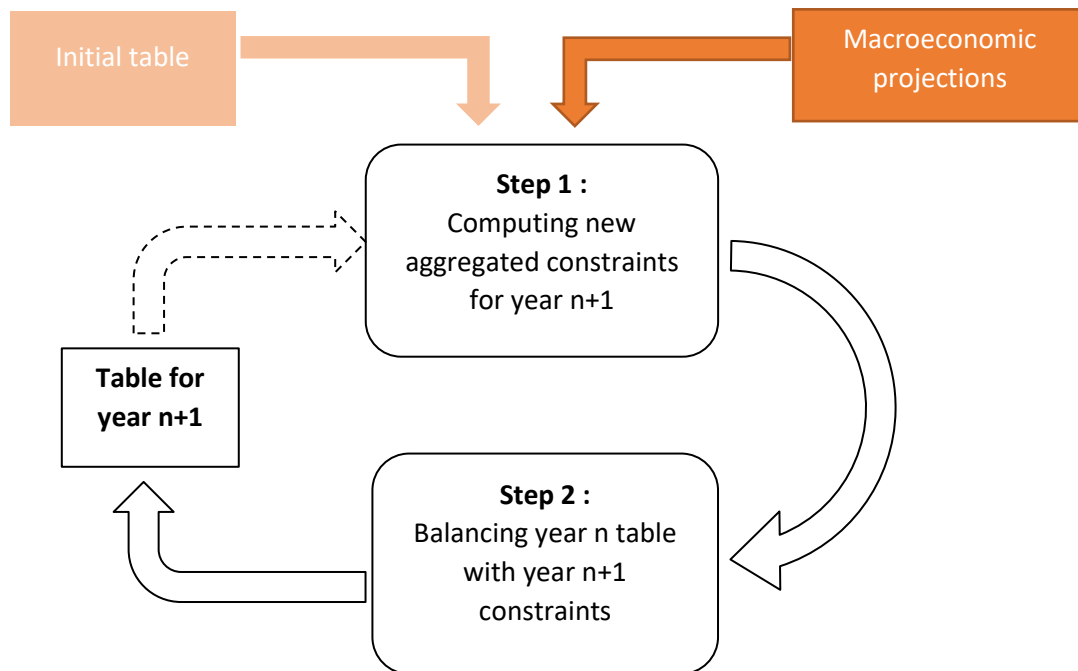


Figure 13 - Global representation of the algorithm

The first step computes constraints as expected demand and available resources for the upcoming year. The second step adapts the previous table to the new constraints, simulating the changes in technologies and in trade relations. The two steps are widely independent, ensuring flexibility and space for further improvements. The first two subsections focus on the main steps of the algorithm. The third subsection discusses some technical issues related to the implementation of the method.

1) Step 1: From macroeconomic projections to aggregated constraints

Our method fits into the SSP framework. The first step of the algorithm transforms a set of economic drivers into aggregated constraints for the matrix to be projected.

a) Linking Primary Inputs and Final Outputs to economic growth

The initial table includes three matrices with subscript 0: trade (T_0), final demand (Y_0), and value-added (V_0). The trade matrix represents the economic flows between different sectors and countries, final demand represents the consumption of final goods and services by different actors in different countries, and value-added gathers the primary inputs for each sector. In the current version of the algorithm, we aggregate final demand and value-added over the different sources of value-added and types of final consumption, so that Y and V are one-dimension arrays.

GDP measures equivalently the global production on a territory, and the revenue generated by the production. The first corresponds therefore to the final demand, and the second is equivalent to the value-added. The evolution of GDP thus directly drives the evolution of both arrays¹.

¹ Technically, this equivalence only holds true if the share of the trade balance in the GDP remains constant over time. Otherwise, the changes in value-added or final demand may flow into international trade. This first version of the method does however not take such phenomenon into account.

Let Y_0 be the final demand array for the initial year and g_i the array associating the cumulated growth projection associated with each country for the year i . We define, for each year i , the following new vector demand Y_i :

$$\forall i, \quad Y_i = Y_0(I + \hat{g}_i) \quad (14)$$

Where $\hat{A}$ represents a diagonalized array. The final demand for year i is therefore simply a scaled-up version of the final demand of year 0. This operation thus assumes there is no change in the relative composition of the final demand. Concerning Value-added, we apply a slightly different equation, because the hypothesis of stability of the trade balance share trigger slight imbalances between final demand and value-added. Consequently, the global change in value-added does not perfectly equalize the global change in final demand. We bridge that gap by applying a uniform corrective factor α_i to the new value-added:

$$\forall i, \quad V_i = V_0 \alpha_i (I + \hat{g}_i) \quad (15)$$

Where:

$$\alpha_i = \frac{|Y_i|}{|V_0 * (I + \hat{g}_i)|} \quad (16)$$

The corrective factor applies uniformly to each sector of the world economy and ensures that the total value-added available at world level meets the total final demand. As for the final demand, the new value-added array is simply scaled-up from the original array. This version of the model does therefore assume a null mobility of the primary inputs: they grow at the same rate in all sectors of the economy.

b) From use and resources to intermediate good assumptions

Previous operation gives us an estimation of the expected final demand and available value-added for each sector for the year to compute. However, the balancing procedure requires converting those indirect output requirements into direct constraints on the trade matrix. To make this conversion, we assume that the technological structure of the projected table will be close to its initial state. Consequently, we consider that the relation between final demand and required input for year n holds true for year $i + 1$. In other words, be L_i the Leontief matrix for year i , we suppose:

$$L_i \sim L_{i+1} \quad (17)$$

Therefore, using equation (III) and the assumption of small technical change, we have the following relation for year $i + 1$, which gives us the required gross output X_{i+1} required to meet the final demand Y_{i+1} :

$$X_{i+1} = L_i Y_{i+1} \quad (18)$$

By subtracting the new final demand array to the gross output, one eventually defines the new inter-industry output O_i . This array defines the volume of intermediate output that each sector should produce in year $i + 1$, or the row-wise sum that will meet the new trade matrix:

$$O_{i+1} = X_{i+1} - Y_{i+1} = (L_i - I)Y_{i+1} \quad (19)$$

Reciprocally, the inter-industry input P_i arises from the equality between gross input and gross output in a balanced MRIOT and represents the volume of intermediate input that each sector needs, or the column-wise sum of the forthcoming trade matrix:

III.2. Method: an algorithm to project future trade relations

$$P_{i+1} = X_{i+1} - V_{i+1} \quad (20)$$

This approach is demand-driven: the evolution of the final demand defines the gross output requirements for each sector. Previously computed value-added simply serve to adjust the input requirement to the gross output, in order to keep the whole matrix balanced. However, we could conversely apply the same logical sequence to compute value-added driven constraints. The main difference with the former approach is that one would need to compute the so-called Ghosh inverse instead of the Leontief inverse. The computation of the Ghosh inverse allows linking an array of primary inputs to the global output, just as the Leontief inverse links the final consumption to the required inputs (Miller and Blair 2009). Both approaches should in theory produce the same output, but practical balancing issues in source data cause minor divergences.

At this point of the algorithm, exogenous GDP projections give use estimations of the future final demand and of the availability of primary inputs for the forthcoming year. Using the Leontief inverse from the previous year, we transform those conditions into expected characteristics for the new trade matrix, in form of row-wise and column-wise requirements. The next step consists in using those characteristics to update the former matrix.

2) Step 2: Updating the matrix

So-called bitproportional technique (or RAS) is a well-known approach to update a MRIO table under exogenous constraints. The idea is to adjust iteratively an original matrix, until it meets new conditions on intermediate inputs and outputs. The algorithm first multiplies every coefficient of each row to meet the row-wise objective value. Then, it multiplies similarly each column to meet the column-wise targets. However, that second operation also changes the row-wise sums. The method therefore consists in repeating those two steps, until the matrix eventually come close enough to the expected characteristics on both directions (Miller and Blair 2009). The usual termination criterion ensures that the difference between each obtained value and the objectives is smaller than a predefined threshold. Numerous authors extended this approach to take in account partial or contradictory data (Lahr and de Mesnard 2004; Lenzen, Gallego, and Wood 2009). However, our approach copes well with a basic RAS algorithm.

This method provides a new table by progressively modifying an original table. Globally, the algorithm results in increasing in priority the coefficients for which the output and the input targets grow the more. Conceptually, one may understand it as increasing the trade from the sectors where the output grows the more to the sectors which increase their need for intermediate inputs, somehow reconciling the changes in offer and demand. However, RAS is a purely mathematical operation, and does not include any economic consideration behind the balancing process. Therefore, it does not take in account the nature of the flows and could reinforce some trade relations without any technological consistence in terms of technological process. Still, this algorithm is well-known, easy to implement, and relatively low in terms of computation requirements. Further work will aim to improve its relevance.

3) Technical Implementation

A few points concerning the implementation of our method are worth mentioning. The first aspect concerns source data. Our method uses a perfectly balanced table as a basis for the computation. However, real table often present imperfect data. For that reason, we apply a preliminary balancing step mainly aiming at balancing the table, so that gross output perfectly equals gross input for every sector. Such balancing requires prioritizing some part of the matrix over others. Practically speaking, constitution of tables relies on surveys and are often less precise for value-added

than for trade or final demand. We therefore re-compute the value added as the difference between gross output and intermediate input, in order to obtain a perfectly balanced matrix.

A second issue concerns the existence of small values in the initial matrix. EORA assign small positive values to all null coefficient for balancing purpose (Lenzen et al. 2012). Most of those coefficients do not represent real flows in the world economy and may worsen the technical consistence of the balancing step, by allowing every single coefficient to grow¹. However, increasing the number of zeros reduces the space of optimization available for the algorithm, which may ultimately impede the convergence of the method. For this last reason, we do not apply a positive threshold to the table, although it could reduce the relevance of the RAS algorithm for simulating technological change.

¹ Since basic RAS algorithms use multiplications to adapt the matrix, they cannot transform null coefficient into non-null economic flows. Therefore RAS algorithms only project changes on already existing coefficients.

III.3. Validation of the performance of the algorithm

III.3. Validation of the performance of the algorithm

Forecasting the exact future of the world economy lies of course far beyond the scope of this paper. However, simulating possible evolutions of global trade requires our model to be able to capture sensible trends in the evolution of international economic flows. We use a back-testing method to check whether our method could reproduce medium terms economic trends. Following verifications are based on EORA 26 time series, available from 1990 to 2015, and on GDP historical data from the World Bank, in current US \$ (World Bank 2019). The following verification methods rely broadly on the same process: we compute time-series starting from different years to 2015 and compare the final table with available EORA data for 2015. We separately test the different steps of our algorithm and conclude with a holistic validation of the full method.

1) First step validation

The first step consists in two consecutive sub-steps: the simulation of expected final demand and value added for the forthcoming year, and the transformation of those arrays into intermediate output constraints via the use of the Leontief inverse.

We test the equivalence between GDP growth and final demand growth by successively applying the growth rate issued from the World Bank to the aggregated final demand of each country. To assess the ability of our model to capture the evolutions of final demand, we capture, for each sector of each country, the following statistics:

$$\theta_n^0 = \frac{\bar{Y}_n - Y_0}{Y_n - Y_0} \quad (XI)$$

Where Y_0 is the aggregated final demand for starting year of estimation, Y_n the reference final demand for year n (from EORA), and $\bar{Y}_n$ the computed final demand for year n . This statistic reflects the part of the change between year 0 and year n captured by the model. We define 1990 as the initial year, and successively compute the final demand for each year until 2015. Figure 14 shows the statistical deviation of the final demand for each year compared to 1990.

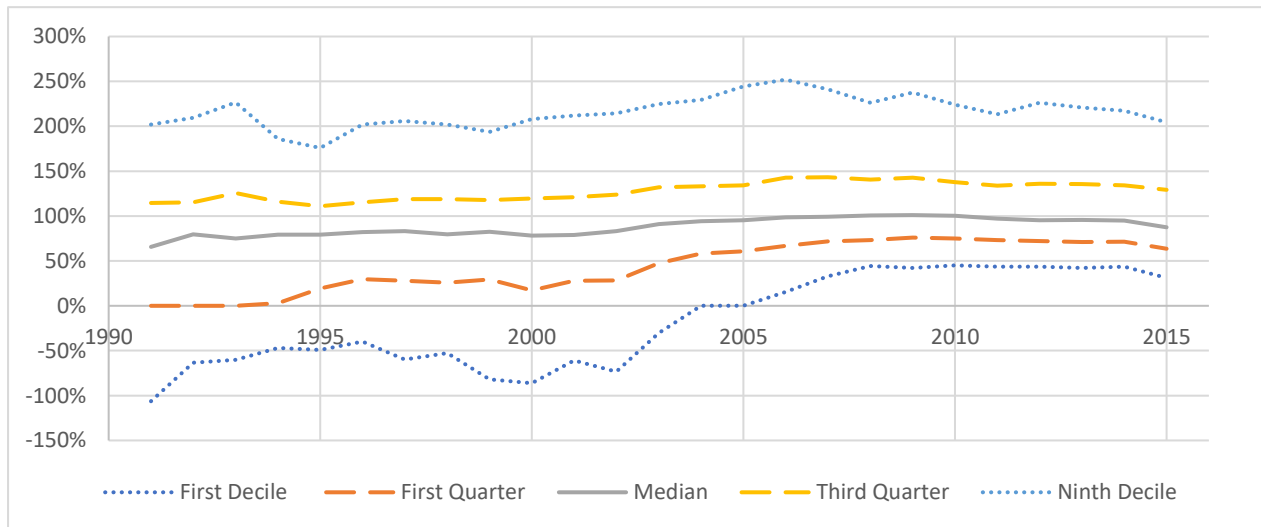


Figure 14 – Evolution of distribution of θ_n^{1990} over time

The first estimation (year 1991) appears only weakly accurate. Median value captures only 61% of the evolution of final demand compared to 1990. Worse, the first quartile is close to 0: a quarter of projected evolutions are in opposition with their observed sense of evolution. However, those inconsistencies do not increase over time, and even slightly reduce. After 11 iterations, median overpass 90%, and the central half of the values fits between 50% and 130% accuracy. Last decile

reaches its top score for 17 iteration, with almost 50% in 2008 and stays globally constant afterward. The last year of simulation is notably less accurate than previous years. Since 2015 is the test table currently available, it may be that this table is less precise than the previous ones. Globally, assumptions on the relation between final demand and GDP evolution seems consistent in a long term. As the relatively strong error in the first rounds highlights, the model seems able capturing the short-term temporary evolutions. However, decrease of deviation over time advocates that the method accurately captures long-term evolution trends of the final demand.

Second sub-step uses the Leontief inverse of the previous year table to compute the gross output required to meet predicted final demand. We test the accuracy of this method by comparing the gross output displayed in Eora for year n to an estimation of this value using the final demand of year n and the Leontief inverse of year $n - 1$. We apply the same statistic as in the last test for the gross output X instead of the final demand Y .

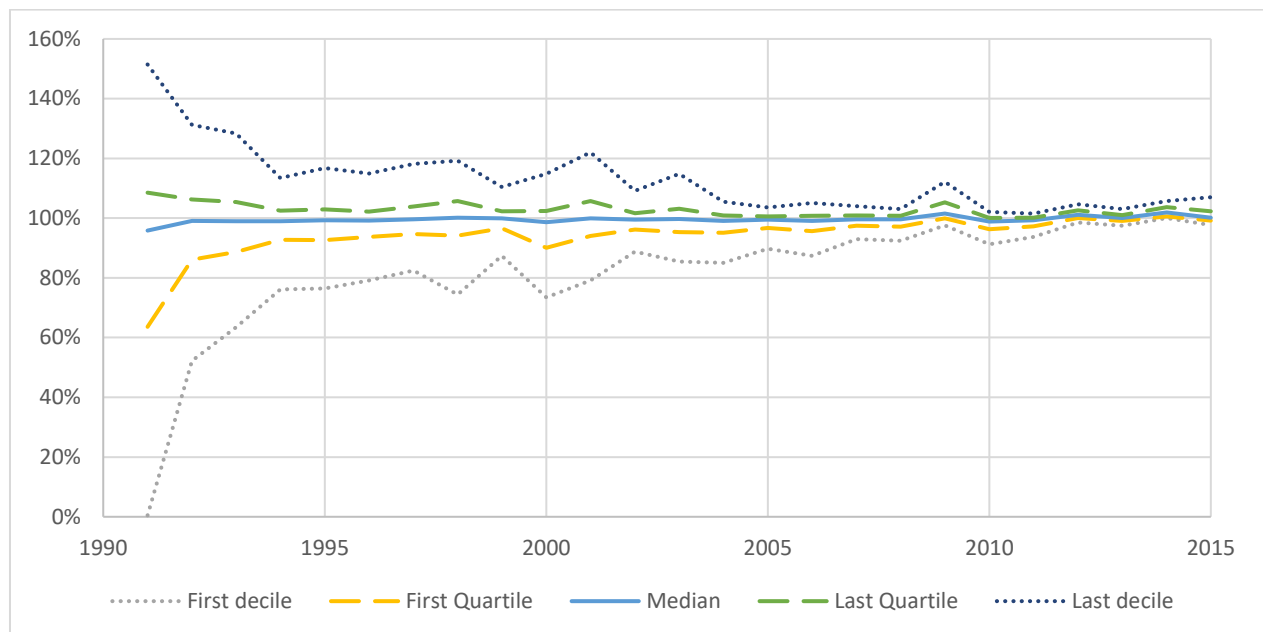


Figure 15 - Distribution of the statistical deviation of the gross output over time

Estimations of gross output using previous year table seems quite accurate. For recent years, more than 80% of values belong to interval [90%-110%] of accuracy. Older years' distribution looks more dispersed, especially in 1991. Collapse of USSR launched a massive restructuration of the world economy that may explain the strong deviation of numerous values. Apart from this period, the assumption of small technological change between year n and year $n + 1$ seems to trigger only a small deviation compared to actual data.

2) Second step validation

The balancing step of the method rely on a RAS algorithm. The ability of this algorithm to capture accurate economic evolutions is still under debate. We test the performance of the RAS algorithm by applying it to the EORA table of year n using the actual constraints of year $n + 1$. We then compare the RAS algorithm to the actual table for year $n + 1$, applying the statistic of equation (XI) to the individual coefficients of the technical matrix.

III.3. Validation of the performance of the algorithm

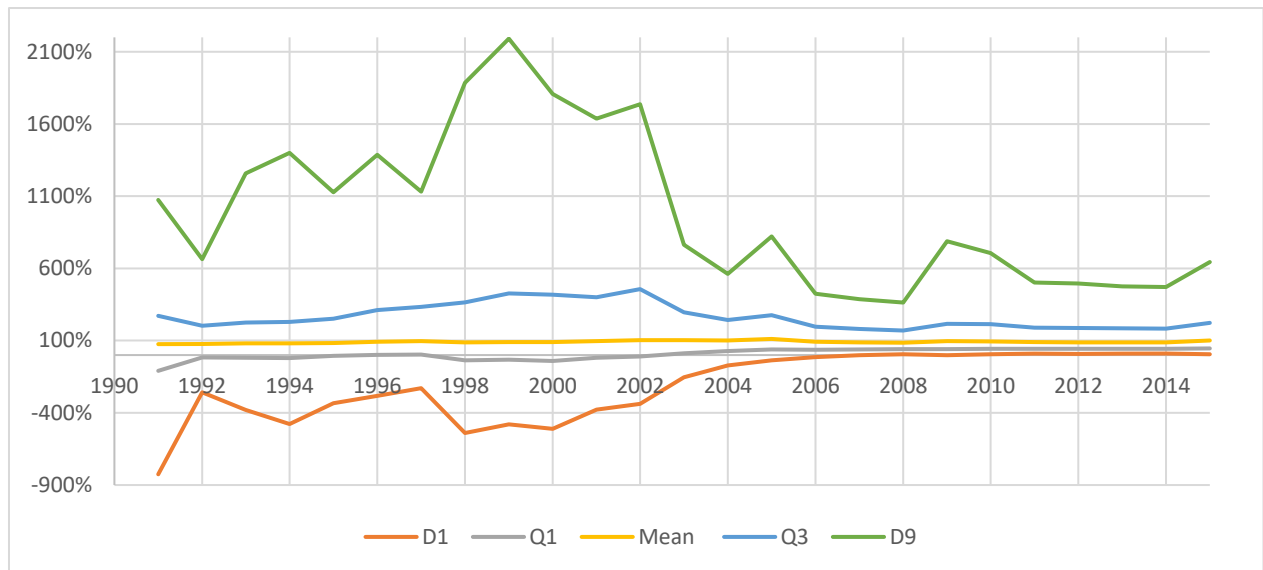


Figure 16 – Distribution of the relative year over year difference between reference data and RAS balancing process

Naïve comparison of evolution of individual coefficients from year to year highlights that RAS algorithm is most of the time not able to capture the evolution of individual coefficient over years. While median values appear close to 100%, many coefficients diverge broadly from reference. First quartile is negative for 10 years out of 25, meaning that more than a quarter estimates evolve in the wrong direction compared to the reference. Surprisingly, there is a clear trend toward convergence over time: value dispersion is much more important before 2003 than after. Still, the top decile and the top quartile remain quite high, respectively above 350 and 190% of deviation. However, EORA contains a lot of small flows, which are close to insignificant for applications of MRIOT. Since the test method does not account for the relative importance of the flows when measuring their deviation, the next figure offers the same analysis applied only to flows higher than 1M US\$.

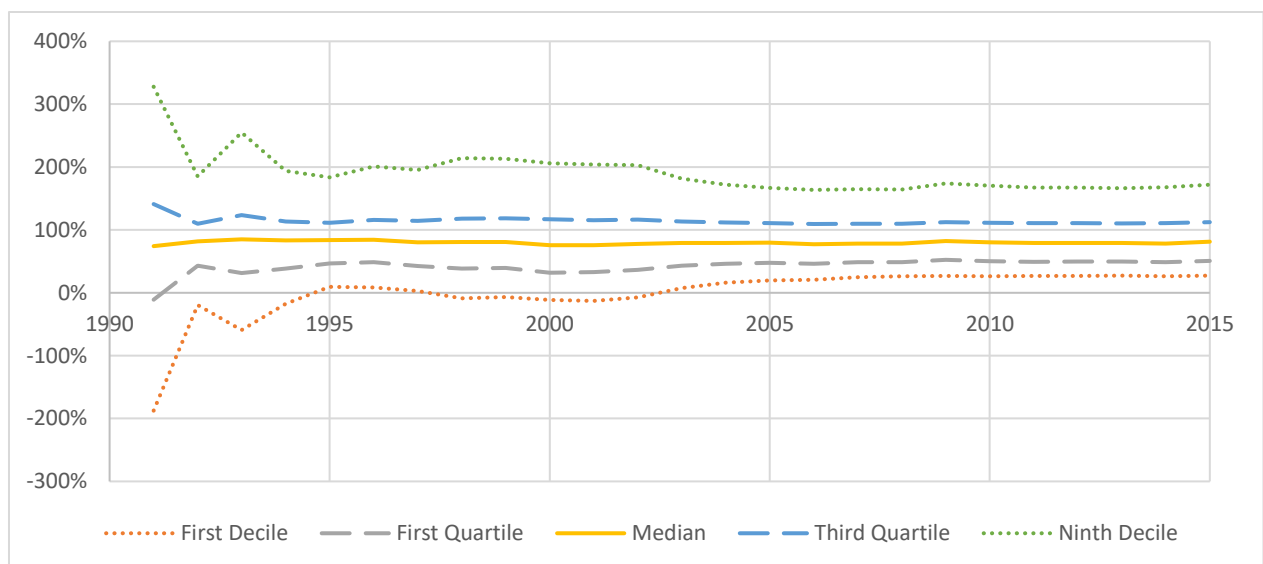


Figure 17 - Distribution of statistical error for flows higher than 1M US\$

When restricting the analysis to major flows, the results become much closer to reference tables and reaches the same order of dispersion than step 1A. After 2003, the RAS algorithm capture between 45% and 180% of the changes for 80% of major coefficients. Interestingly, median is slightly lower for major coefficients compared to all coefficient (around 80%). It could mean that the process underestimates some major coefficients to the benefit of small coefficients. Previous application of a

threshold to EORA tables could help tackling this issue. However, ignoring small coefficient may trigger numerical issues, by impeding the computation of the Leontief inverse and by constraining too much the balancing process.

3) Holistic validation

Independent tests on the different steps of the method reveal limited evidences on the ability of our method to capture evolution trends of the global economy, because we don't know how the errors appearing at the different steps of the algorithm propagate over iterations. We therefore apply the full method to build a 14-year time series, from 2000 to 2014, and we compare the quality of the estimation compared to independently applied steps.

Since step two balances the matrix to meet the aggregated constraints defined through step one, step 1A and 1B completely determine evolution of gross output and final demand over time. Under the assumption that step two converges successfully, assessing the consistency of the aggregated output throughout the process is no different to assessing its consistency through step one.

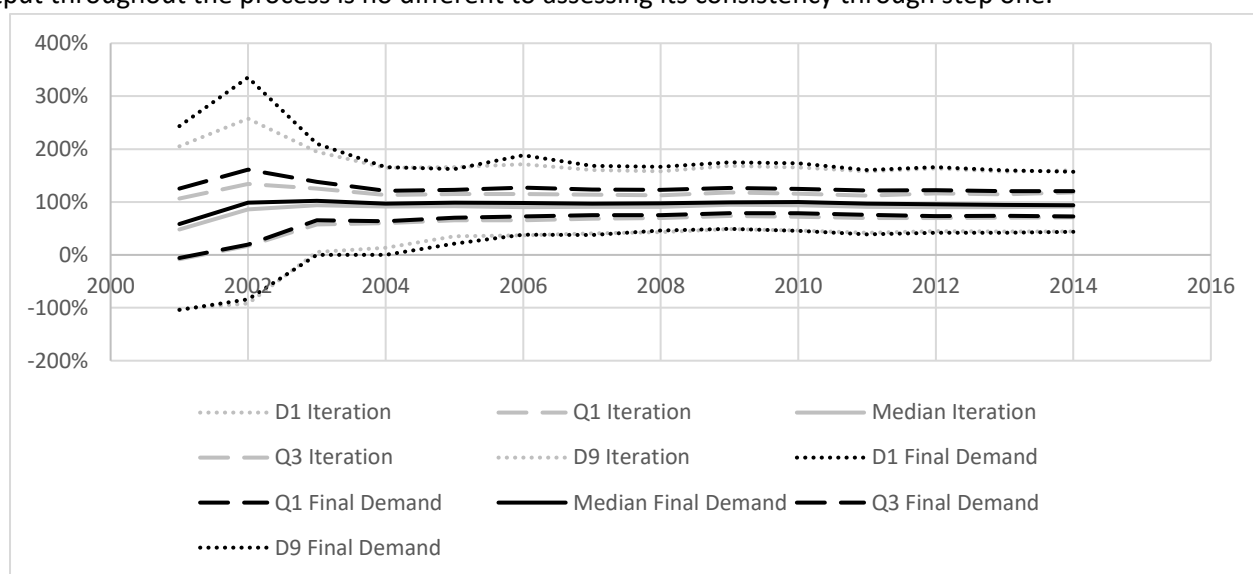


Figure 18 – Distribution of statistic deviation of gross output in reference to starting year, comparison with estimation of Final Demand over years.

On the middle term, iterations of the process over time trigger deviations between the actual and the computed final demand which are distributed likewise the errors caused by step 1B (see Figure 14). The deviation of the final demand and of the gross output are very close to each other. Arguably, the transformation of final demand constraints into gross output constraints creates much smaller deviations than the estimation of the final demand. In other words, deviation induced by step 1B is an order of magnitude above deviation caused by step 1A, and the assumption of small technological change holds better than the uniform growth of consumption at the rate of GDP.

Also, this figure underlines that our method does not diverge over time. After a few years, deviation regarding the starting year appears stable, and quite limited: 80% of the values estimates the evolution of final demand between 50% and 150% of their actual evolution. The broader dispersion of the error in the two first years highlights that the algorithm captures poorly short-term evolution compared to long term trends. A first piece of explanation is that yearly economic data may suffer more erratic changes, which smooth at longer term. A second and related piece of evidence is the fact that volumes grow over time, mathematically diminishing the relative deviation.

Finally, Figure 19 shows the distribution of the deviation for individual coefficients simulated from 2000 to 2015 using the full method.

III.3. Validation of the performance of the algorithm

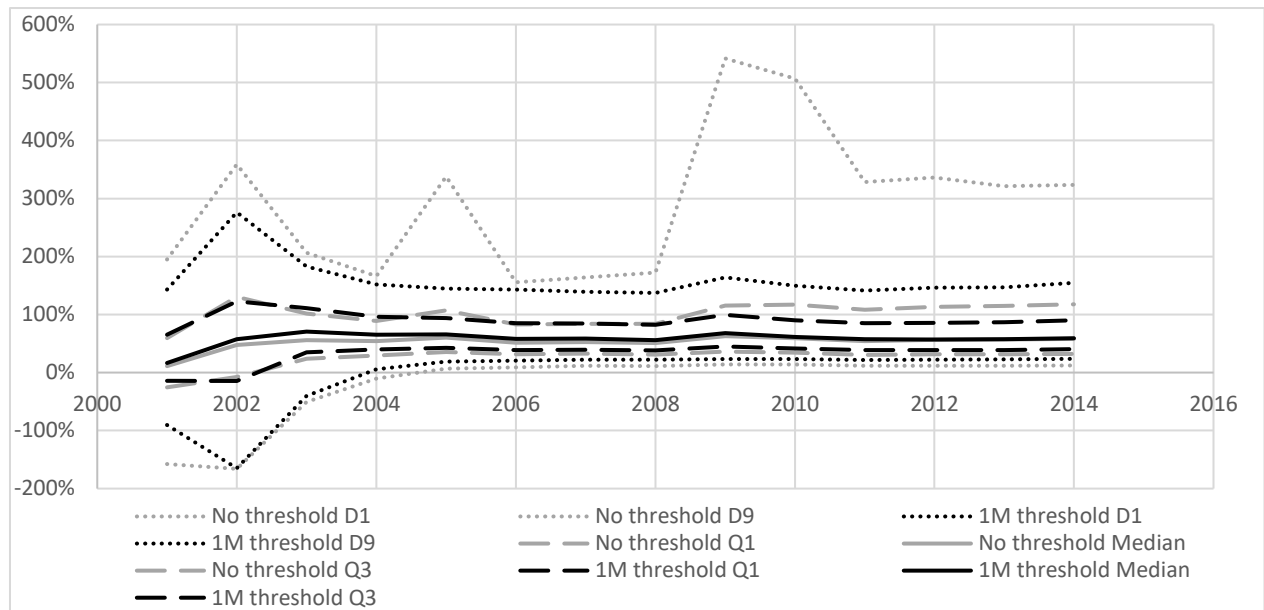


Figure 19 –Distribution of the relative evolution of individual coefficients captured by the model since 2000. Dark lines refer to flows bigger than 1M US\$

At middle term, the coefficients provided by our method appear often underestimated, but in an acceptable range of deviation. Since 2004, more than 90% of big flows estimations capture the good sense of evolution of the flows. As reflects the wider inter-decile range for all coefficient, small coefficients capture relatively more change than coefficients higher than 1M US\$. First quartile and decile appear close to the median, which is closer to 50% than 100%. Since third quartile lays close to 100%, around 75% of coefficients are underestimated by the model. Interestingly, there is no obvious divergence of the model over time. The algorithm appears producing stable estimations of the evolution of commercial flows, and 80% of the major flows belong to the range [10%-150%] after a 15 years simulation. While this range remains quite large, it appears however quite satisfactory considering the simplicity of our model and the complexity of the problem.

The main insights of the performance assessment of the method are the following. First, individual steps proceed unequally well. Step 1B gives obviously more robust results than step 1A and step 2. There is therefore a substantial room for improvement for those two parts. Secondly, short term changes in the economy appear to messy to be captured by the current model. All tests agree on the fact that deviation between the simulations and the reference database flatten over time. Arguably, the method performs better for middle-term trends. Lastly, both the succession of steps and iterations appears relatively stable. Holistic test shows no exponential deviations due to the succession of operations, but rather constant relative deviation. This last feature is highly encouraging, and the improvement of the independent step of the algorithm may further improve the global accuracy of the process.

III.4. Discussion & further work

We introduce a new method to simulate MRIO time series under exogenous conditions. The iterative process relies on two highly largely independent steps, creating first new constraints using GDP projections, and then using those constraints to extrapolate a past table. The algorithm relies on well-established theoretical basis, requires a limited number of assumptions, and is relatively easy to implement comparatively to others modelling approaches (Integrated Assessment Modelling, Computable General Equilibrium). Back testing of the algorithm strongly supports the *better than persistence* hypothesis. While individual values may diverge largely from observation, our model arguably captures global trends of the world economy over middle term. It also performs interestingly well in terms of time consistency, triggering no obvious exponential divergence over time. Moreover, the modularity of our approach may support further improvements.

The two main sources of deviation lay for the same order of magnitude in the forecasting of final demand from GDP growth rate and of the application of the RAS algorithm for balancing purpose. The use of GDP as a direct driver for final demand blinks the evolution of national trade balance and does not account for the evolution of the composition of the final demand. Implementation of assumptions on the global evolution of international could both improve our estimation and provide new possibility for scenario making. At the same time, introducing flexibility in the final demand's composition or in primary input allocation could open new possibilities of the method, but at the cost of more restrictive assumptions. Concerning the balancing step, RAS is a fundamentally mathematical approach, and tends to distort the detailed structure of the world economy. For instance, it overestimates the evolution of small flows. While this has a limited effect on aggregated measures, such distortion may limit the application of estimated tables for detailed analysis. Further work may therefore aim constraining the RAS algorithm in order to improve its accuracy.

The next steps of this research embody several perspectives. First, there is a need to find more precise ways for assessing the performances of the current method. This report provides an analysis in terms of error distribution, which does not reflect the weight of those errors on global figures. Therefore, the method appears quantitatively encouraging, but we miss a protocol to assess more qualitative aspects, such as the credibility of the table in terms of technological changes. A main challenge on this path is the size of the output file. Advanced data analysis tools could obviously provide relevant insights on this issue. Secondly, there is plenty of room for improving steps 1A and B. In both cases, the introduction of more advanced methods could increase both the accuracy of the modelling and open new possibilities for scenario customization. Such improvement could either be analytical or statistical.

Eventually, this method highlights the potential of IO tables for prospective application. The rather simple model developed for the purpose of this report appears as a promising way toward quantitative insights on the future international trade. With no doubt, such forecasted table could be of great use for assessing the potential impact of climate change on supply chains. Such insights would be of great use for numerous application cases, from public policies to insurance policies or stakeholder's transparency.

CONCLUSION

MRIO is a macro-economic approach aiming at capturing the systemic linkages throughout transnational economic networks. Thanks to its completeness and its consistency, it supports sound mathematical properties, allowing a direct assessment of direct and indirect inputs requirements of any industrial output. The extension of MRIO to social and environmental databases allows the translation of those economic requirements into material externalities. Such model permits capturing an immediate snapshot of the broad social and environmental externalities of any monetarized interaction between economic actors. Such top-down assessment however encounters limits in numerical and sectorial precision. However, EEIO techniques offer powerful insights on the global trends and on the embodiment of environmental externalities throughout international trade.

The tool developed for the account of PwC is a direct application of those principles. The operationalization of such technique however triggers supplementary questions in terms of data reliability and compatibility with other approaches. A critique examination of those aspects highlights that EEIO is not a silver bullet and suffers undeniable numerical uncertainties. However, this technique provides valuable insights on system-wide issues, while being flexible and easy to use. By providing quantitative assessments on supply chains with an unmatched efficiency, such tool meets adequately the growing need for rational approaches to sustainability issues.

The last chapter of this report highlights the high potential of MRIO for prospective analysis. Such new field of application fits the global need for quantitative scenarios to assess the challenges society will have to tackle in the future. The method thereby introduced offers a promising path toward forecasted MRIO tables. Such table could support a wide range of analysis, from the assessment of climate events on the world economy, to the strategic impacts of some transformation pathways. However, this work is still under progress. First, because the validation method does not offer detailed insights on the qualitative relevance of the method. Secondly, there is space for improving each step of the algorithm, in order to enhance the parametrization of the simulations and to reflect more accurately qualitative differences between development scenarios.

As a conclusion, this report offers an overview on current and promising applications of MRIO tables. It appears that the MRIO framework draws its main force from its completeness and theoretical consistence. The mathematical implications of such qualities make MRIO a favored tool for studying the systemic interactions in the world economy. Both the complexification of supply chains and the growing incentives to understand the risks and externalities caused by human activities advocate for an in-depth understanding of the high-order phenomenon appearing in supply chains. MRIO-based tools appear already as key approaches to tackle such issues at any scale, from operational decision-making to longer term strategic decision assessment. While recognizing the limits of such top-down approach, MRIO has definitively the potential to deliver by now critical insights for policy-makers, as underlines the trend toward the appropriation of such technology by economic actors. Simultaneously, MRIO undergoes a living research effort, which promises delivering further applications cases of this tool, including for prospective analysis.

In a broader perspective, this internship was the occasion of an immersion at the interface of academic and professional universes. The focus of my work highlights the value of the interactions between economic actors and researchers. The transfer of state-of-art academic techniques to businesses is a key game-changer, opening new business opportunities for strategic decision-making. However, such valorization is none of an easy process. Academic research requires the acknowledgement of the limits and uncertainties associated with any technique, while economic decision requires definite and reliable insights. The answer to whether a result is faithful or not is

crucial for decision making but corresponds to a blurred notion in academic research. The operationalization of academic techniques therefore requires setting arbitrarily the conditions under which a result hold and make those conditions clear for a relevant valorization of the results.

I consider the outcomes of my internship as positive, in consideration to its initial targets. I was able to deliver a fully operational tool for the client, using state-of-art techniques of economic modelling. I believe this tool provides a sensible approach to tackle existing but also forthcoming business priorities of PwC's clients. I think these techniques fill a gap in the current technological toolbox of the firm and will help the company to maintain its leading position on the field of businesses' sustainability advisory. I also cared to design the tool to be customizable and improvable, to ensure its durability. Simultaneously, my stay at PIK allowed me to produce promising results on two relevant topics. First, I integrated new indicators in a widely used database and applied it to the examination of the issue of global GHG emissions. Those practical outcomes have a strong interest for a better understanding of the footprint of consumption patterns and may support further applications. Then, I sketched a method for forecasting MRIO tables along SSPs scenarios, paving the way for high potential prospective applications of MRIO. This method provides promising results and offers plenty of room for further improvement.

Those outcomes suit clear perspectives. PwC's tool has the potential to address a wide range of operational problem in the future years. The design of the tool and the joined documentation will also permit the adaptation of the tool to further developments of the databases at a minimal cost. Concerning PIK's research, the outcomes of the internship provide a basis for the preparation of a doctorate thesis under the supervision of Leonie Wenz and Anders Levermann. In a shorter term, we plan to submit academic publications based on the two main research outcomes of this thesis.

ABBREVIATIONS

CBA	Consumption Based Accounting
EEIO	Environmentally Extended Input Output
EWMFA	Economy-Wide Material Flow Analysis
FTE	Full Time Equivalent
GDP	Gross Domestic Product
Gg	Gigagrams (or kilotons)
GHG	Greenhouse Gases
ILO	International Labor Organization
IO	Input Output
IPCC	International Panel on Climate Change
LCA	Life Cycle Analysis
LCI	Life Cycle Inventory
IEA	International Energy Agency
MRIO	Multi Regional Input Output
MRIOT	Multi Regional Input Output Table
PBA	Process Based Accounting (with LCA)
PBA	Production Based Accounting (for footprint accounting)
PIK	Potsdam Institute for Climate Impact Research
SSP	Shared Socioeconomic Pathways
SUT	Supply Use Table
Toe	Ton of oil equivalent ($= 4.187 \cdot 10^{16} J$)
UNEP	United Nations Environment Program
UNFCC	United Nations Framework Convention on Climate Change
WFN	Water Footprint Network

GLOSSARY

Basic Price	Or producer price. It is the price of a commodity as collected by its producer. It does not include any tax or subsidy but may include the margin.
CO2e	Unit of radiative forcing increase due to the emissions of greenhouse gases, expressed in carbon dioxide equivalently emitted. The conversion of gases into CO2e relies on their Global Warming Potential, regularly updated by the IPCC on the light of the more recent information available ¹ .
Commodities	Designates standard goods and services. Each item of a commodity is supposed to be economically equivalent to every other item of the same commodity.
Direct Externalities	Designates all externalities directly caused by a process.
Economic sector	Classification of activity by position in supply chains or kind of production.
Externalities	In economics, externalities designate positive, neutral or negative influences affecting subjects who did not cause them.
EEIO	Input Output table that includes environmental indicators.
EORA	World EEIO database supporting the tool. It is available in two sectorial definition: Eora 26 and Full Eora.
EORA 26	Simplified version of the Full Eora database. Provide a homogenous sectorial classification of 26 sectors. Slightly less precise than Full Eora.
Final goods	Designates commodities that will not undergo further transformation but are to be used for final consumption.
Final Price	Or consumer price. It is the public price of a commodity. It includes the basic price and several layers, including taxes, subsidies, transportation... It may differ in function of the type of consumer
Footprint	Inventory of the externalities associated with some consumption, including the externalities embodied in previous steps of the life cycle of the products.
FTE	Measure the number of workers in a country as if everyone worked in full time. Two half-time workers count therefore as only one Full Time Equivalent.
Full Eora	Main table of the EORA database. Provides the most detailed sectorial definition for each country.
Kyoto GHG	Weighted basket of greenhouse gases, defined by the IPCC.
Indirect externalities	Qualifies the externalities required to produce all primary and intermediate goods that ultimately serve to the production of a final good or service. Those externalities issue from previous (and sometimes posteriors) process of the value chain.

¹http://www.ghgprotocol.org/sites/default/files/ghgp/Global-Warming-Potential-Values%20%28Feb%2016%202016%29_1.pdf

Intermediary Inputs	Input coming from another sector of the economy, by opposition to primary inputs.
Input Output Model	Simplified representation of an economy as a two-dimensional square table, where each coefficient measures the output from one sector serving as an input for another sector.
Life cycle analysis	Assessment of the environmental externalities associated with all the stages of existence of a product, from material inputs extraction to its disposal. ISO standards 14040 and 14044 detail its objective and define its four phases: Goal and scope definition, Inventory Analysis, Impact Assessment and Interpretation.
Life cycle inventory	One of the main phases of the LCA as defined by standard ISO 14040 and 14044. Consists in modelling all materials and externalities flowing in and out the system boundaries.
MRIO table	Input Output table picturing numerous regions.
Primary inputs	Inputs not provided by another economic sector, but by consumption of labor and capital.
RAS algorithm	Mathematical operation transforming a basic matrix to meet new requirements in terms of row-wise and column-wise sums.
Satellite Matrix	Part of a MRIO table containing non-economic data (for instance socio-environmental indicators).
Sectorial disaggregation/classification/definition	List of economic sectors used to model the economic activity of a country. Higher sectorial disaggregation gives a more accurate representation of an economy but require more data processing.
Technical matrix	Normalized form of the trade matrix. Coefficients represent the share of one sector's output in another sector required input mix.
Territorial Externalities	Externalities taking place in a defined territory. Also called local externalities.
Trade matrix	Main matrix of a MRIO table. Represent all interactions between sectors.
Water (Blue)	Freshwater, coming both from surface and from ground.
Water (Green)	Precipitation water stored in soil and in vegetation.
Water (Grey)	Required freshwater volume to dilute water pollutant enough to reach water quality standards

BIBLIOGRAPHY

- Akizu-Gardoki, Ortzi, Gorka Bueno, Thomas Wiedmann, Jose Manuel Lopez-Guede, Iñaki Arto, Patxi Hernandez, and Daniel Moran. 2018. "Decoupling between Human Development and Energy Consumption within Footprint Accounts." *Journal of Cleaner Production* 202 (November): 1145–57. <https://doi.org/10.1016/j.jclepro.2018.08.235>.
- Alsamawi, Ali, Joy Murray, Manfred Lenzen, Daniel Moran, and Keiichiro Kanemoto. 2014. "The Inequality Footprints of Nations: A Novel Approach to Quantitative Accounting of Income Inequality." Edited by Rodrigo Huerta-Quintanilla. *PLoS ONE* 9 (10): e110881. <https://doi.org/10.1371/journal.pone.0110881>.
- Aslam, Aqib, Natalija Novta, and Fabiano Rodrigues-Bastos. 2018. "Calculating Trade in Value Added." Working Paper. IMF Working Paper.
- Benoit-Norris, Catherine, Deana Aulisio Cavan, and Gregory Norris. 2012. "Identifying Social Impacts in Product Supply Chains: Overview and Application of the Social Hotspot Database." *Sustainability* 4 (9): 1946–65. <https://doi.org/10.3390/su4091946>.
- Box, George E. P., and Norman R. Draper. 1987. *Wiley Series in Probability and Mathematical Statistics. Empirical Model-Building and Response Surfaces*. John Wiley & Sons. Oxford, England.
- Brizga, Janis, Kuishuang Feng, and Klaus Hubacek. 2017. "Household Carbon Footprints in the Baltic States: A Global Multi-Regional Input–Output Analysis from 1995 to 2011." *Applied Energy* 189 (March): 780–88. <https://doi.org/10.1016/j.apenergy.2016.01.102>.
- Crawford, Robert H., Paul-Antoine Bontinck, André Stephan, Thomas Wiedmann, and Man Yu. 2018. "Hybrid Life Cycle Inventory Methods – A Review." *Journal of Cleaner Production* 172 (January): 1273–88. <https://doi.org/10.1016/j.jclepro.2017.10.176>.
- Crespo Cuaresma, Jesús. 2017. "Income Projections for Climate Change Research: A Framework Based on Human Capital Dynamics." *Global Environmental Change* 42 (January): 226–36. <https://doi.org/10.1016/j.gloenvcha.2015.02.012>.
- Dellink, Rob, Jean Chateau, Elisa Lanzi, and Bertrand Magné. 2017. "Long-Term Economic Growth Projections in the Shared Socioeconomic Pathways." *Global Environmental Change* 42 (January): 200–214. <https://doi.org/10.1016/j.gloenvcha.2015.06.004>.
- Distelkamp, Martin, and Mark Meyer. 2019. "Pathways to a Resource-Efficient and Low-Carbon Europe." *Ecological Economics* 155 (January): 88–104. <https://doi.org/10.1016/j.ecolecon.2017.07.014>.
- Eggleston, Simon, Leandro Buendia, Kyoko Miwa, Todd Ngara, and Kiyoto Tanabe. 2006. "2006 Guidelines for National Greenhouse Gas Inventories." IPCC.
- Fischer-Kowalski, M., F. Krausmann, S. Giljum, S. Lutter, A. Mayer, S. Bringezu, Y. Moriguchi, H. Schütz, H. Schandl, and H. Weisz. 2011. "Methodology and Indicators of Economy-Wide Material Flow Accounting: State of the Art and Reliability Across Sources." *Journal of Industrial Ecology* 15 (6): 855–76. <https://doi.org/10.1111/j.1530-9290.2011.00366.x>.
- Global Carbon Project. 2018. "Supplemental Data of Global Carbon Project 2018." Global Carbon Project. <https://doi.org/10.18160/gcp-2018>.
- "Global Material Flows Database." 2018. UNEP. <https://www.resourcepanel.org/global-material-flows-database>.
- Gütschow, Johannes, Louise Jeffery, and Robert Gieseke. 2019. "The PRIMAP-Hist National Historical Emissions Time Series (1850-2016)." GFZ Data Services. <https://doi.org/10.5880/pik.2019.001>.
- Haraguchi, Masahiko, and Upmanu Lall. 2015. "Flood Risks and Impacts: A Case Study of Thailand's Floods in 2011 and Research Questions for Supply Chain Decision Making." *International Journal of Disaster Risk Reduction* 14 (December): 256–72. <https://doi.org/10.1016/j.ijdrr.2014.09.005>.
- IPCC. 2012. "Managing the Risks of Extreme Events and Disasters to Advance Climate Change Adaptation." Special Report. IPCC.

- Islam, Samantha, S.G. Ponnambalam, and Hon Loong Lam. 2016. "Review on Life Cycle Inventory: Methods, Examples and Applications." *Journal of Cleaner Production* 136 (November): 266–78. <https://doi.org/10.1016/j.jclepro.2016.05.144>.
- Janssens-Maenhout, Greet, Monica Crippa, Diego Guizzardi, Marilena Muntean, Edwin Schaaf, Frank Dentener, Peter Bergamaschi, et al. 2017. "EDGAR v4.3.2 Global Atlas of the Three Major Greenhouse Gas Emissions for the Period 1970–2012." *Earth System Science Data Discussions*, August, 1–55. <https://doi.org/10.5194/essd-2017-79>.
- Jiborn, Magnus, Astrid Kander, Viktoras Kulionis, Hana Nielsen, and Daniel D. Moran. 2018. "Decoupling or Delusion? Measuring Emissions Displacement in Foreign Trade." *Global Environmental Change* 49 (March): 27–34. <https://doi.org/10.1016/j.gloenvcha.2017.12.006>.
- Kander, Astrid, Magnus Jiborn, Daniel D. Moran, and Thomas O. Wiedmann. 2015. "National Greenhouse-Gas Accounting for Effective Climate Policy on International Trade." *Nature Climate Change* 5 (5): 431–35. <https://doi.org/10.1038/nclimate2555>.
- Kanemoto, K., D. Moran, M. Lenzen, and A. Geschke. 2014. "International Trade Undermines National Emission Reduction Targets: New Evidence from Air Pollution." *Global Environmental Change* 24 (January): 52–59. <https://doi.org/10.1016/j.gloenvcha.2013.09.008>.
- Kanemoto, Keiichiro, Manfred Lenzen, Glen P. Peters, Daniel D. Moran, and Arne Geschke. 2012. "Frameworks for Comparing Emissions Associated with Production, Consumption, And International Trade." *Environmental Science & Technology* 46 (1): 172–79. <https://doi.org/10.1021/es202239t>.
- Kc, Samir, and Wolfgang Lutz. 2017. "The Human Core of the Shared Socioeconomic Pathways: Population Scenarios by Age, Sex and Level of Education for All Countries to 2100." *Global Environmental Change* 42 (January): 181–92. <https://doi.org/10.1016/j.gloenvcha.2014.06.004>.
- Kennelly, C., M. Berners-Lee, and C.N. Hewitt. 2019. "Hybrid Life-Cycle Assessment for Robust, Best-Practice Carbon Accounting." *Journal of Cleaner Production* 208 (January): 35–43. <https://doi.org/10.1016/j.jclepro.2018.09.231>.
- Koning, Arjan de, Martin Bruckner, Stephan Lutter, Richard Wood, Konstantin Stadler, and Arnold Tukker. 2015. "Effect of Aggregation and Disaggregation on Embodied Material Use of Products in Input–Output Analysis." *Ecological Economics* 116 (August): 289–99. <https://doi.org/10.1016/j.ecolecon.2015.05.008>.
- Lahr, Michael, and Louis de Mesnard. 2004. "Biproportional Techniques in Input-Output Analysis: Table Updating and Structural Analysis." *Economic Systems Research* 16 (2): 115–34. <https://doi.org/10.1080/0953531042000219259>.
- Le Quéré, Corinne, Robbie M. Andrew, Pierre Friedlingstein, Stephen Sitch, Judith Hauck, Julia Pongratz, Penelope A. Pickers, et al. 2018. "Global Carbon Budget 2018." *Earth System Science Data* 10 (4): 2141–94. <https://doi.org/10.5194/essd-10-2141-2018>.
- Leimbach, Marian, Elmar Kriegler, Niklas Roming, and Jana Schwanitz. 2017. "Future Growth Patterns of World Regions – A GDP Scenario Approach." *Global Environmental Change* 42 (January): 215–25. <https://doi.org/10.1016/j.gloenvcha.2015.02.005>.
- Lenzen, Manfred. 2000. "Errors in Conventional and Input-Output–Based Life–Cycle Inventories." *Journal of Industrial Ecology* 4 (4): 127–48. <https://doi.org/10.1162/10881980052541981>.
- Lenzen, Manfred, Blanca Gallego, and Richard Wood. 2009. "MATRIX BALANCING UNDER CONFLICTING INFORMATION." *Economic Systems Research* 21 (1): 23–44. <https://doi.org/10.1080/09535310802688661>.
- Lenzen, Manfred, Keiichiro Kanemoto, Daniel Moran, and Arne Geschke. 2012. "Mapping the Structure of the World Economy." *Environmental Science & Technology* 46 (15): 8374–81. <https://doi.org/10.1021/es300171x>.
- Lenzen, Manfred, Daniel Moran, Anik Bhaduri, Keiichiro Kanemoto, Maksud Bekchanov, Arne Geschke, and Barney Foran. 2013. "International Trade of Scarce Water." *Ecological Economics* 94 (October): 78–85. <https://doi.org/10.1016/j.ecolecon.2013.06.018>.

Bibliography

- Leontief, Wassily W. 1936. "Quantitative Input and Output Relations in the Economic Systems of the United States." *The Review of Economics and Statistics* 18 (3): 105. <https://doi.org/10.2307/1927837>.
- Levermann, Anders. 2014. "Climate Economics: Make Supply Chains Climate-Smart." *Nature* 506 (7486): 27–29. <https://doi.org/10.1038/506027a>.
- Malik, Arunima, Manfred Lenzen, Peter J. Ralph, and Bojan Tamburic. 2015. "Hybrid Life-Cycle Assessment of Algal Biofuel Production." *Bioresource Technology* 184 (May): 436–43. <https://doi.org/10.1016/j.biortech.2014.10.132>.
- Mercure, Jean-Francois, Hector Pollitt, Neil R. Edwards, Philip B. Holden, Unnada Chewpreecha, Pablo Salas, Aileen Lam, Florian Knobloch, and Jorge E. Vinuales. 2018. "Environmental Impact Assessment for Climate Change Policy with the Simulation-Based Integrated Assessment Model E3ME-FTT-GENIE." *Energy Strategy Reviews* 20 (April): 195–208. <https://doi.org/10.1016/j.esr.2018.03.003>.
- Meyer, Bernd, and Gerd Ahlert. 2019. "Imperfect Markets and the Properties of Macro-Economic-Environmental Models as Tools for Policy Evaluation." *Ecological Economics* 155 (January): 80–87. <https://doi.org/10.1016/j.ecolecon.2017.06.017>.
- Miller, Ronald E, and Peter D Blair. 2009. "Input–Output Analysis: Foundations and Extensions, Second Edition," Cambridge University Press, , 784.
- Moran, Daniel, and Keiichiro Kanemoto. 2017. "Identifying Species Threat Hotspots from Global Supply Chains." *Nature Ecology & Evolution* 1 (1): 0023. <https://doi.org/10.1038/s41559-016-0023>.
- Moran, Daniel, and Richard Wood. 2014. "CONVERGENCE BETWEEN THE EORA, WIOD, EXIOBASE, AND OPENEU'S CONSUMPTION-BASED CARBON ACCOUNTS." *Economic Systems Research* 26 (3): 245–61. <https://doi.org/10.1080/09535314.2014.935298>.
- Oita, Azusa, Arunima Malik, Keiichiro Kanemoto, Arne Geschke, Shota Nishijima, and Manfred Lenzen. 2016. "Substantial Nitrogen Pollution Embedded in International Trade." *Nature Geoscience* 9 (2): 111–15. <https://doi.org/10.1038/ngeo2635>.
- Okuyama, Yasuhide, and Joost R. Santos. 2014. "DISASTER IMPACT AND INPUT–OUTPUT ANALYSIS." *Economic Systems Research* 26 (1): 1–12. <https://doi.org/10.1080/09535314.2013.871505>.
- O'Neill, Brian C., Elmar Kriegler, Kristie L. Ebi, Eric Kemp-Benedict, Keywan Riahi, Dale S. Rothman, Bas J. van Ruijven, et al. 2017. "The Roads Ahead: Narratives for Shared Socioeconomic Pathways Describing World Futures in the 21st Century." *Global Environmental Change* 42 (January): 169–80. <https://doi.org/10.1016/j.gloenvcha.2015.01.004>.
- Otto, C., S.N. Willner, L. Wenz, K. Frierler, and A. Levermann. 2017. "Modeling Loss-Propagation in the Global Supply Network: The Dynamic Agent-Based Model Acclimate." *Journal of Economic Dynamics and Control* 83 (October): 232–69. <https://doi.org/10.1016/j.jedc.2017.08.001>.
- Owen, Anne, Kjartan Steen-Olsen, John Barrett, Thomas Wiedmann, and Manfred Lenzen. 2014. "A STRUCTURAL DECOMPOSITION APPROACH TO COMPARING MRIO DATABASES." *Economic Systems Research* 26 (3): 262–83. <https://doi.org/10.1080/09535314.2014.935299>.
- Pomponi, Francesco, and Manfred Lenzen. 2018. "Hybrid Life Cycle Assessment (LCA) Will Likely Yield More Accurate Results than Process-Based LCA." *Journal of Cleaner Production* 176 (March): 210–15. <https://doi.org/10.1016/j.jclepro.2017.12.119>.
- "Population, Total." 2019. *World Bank DataBank* (blog). 2019. <https://donnees.banquemondiale.org/indicateur/SP.POP.TOTL?end=2015&start=1960>.
- Potter, Philip, Navin Ramankutty, Elena M. Bennett, and Simon D. Donner. 2010. "Characterizing the Spatial Patterns of Global Fertilizer Application and Manure Production." *Earth Interactions* 14 (2): 1–22. <https://doi.org/10.1175/2009EI288.1>.
- Riahi, Keywan, Detlef P. van Vuuren, Elmar Kriegler, Jae Edmonds, Brian C. O'Neill, Shinichiro Fujimori, Nico Bauer, et al. 2017. "The Shared Socioeconomic Pathways and Their Energy, Land Use, and Greenhouse Gas Emissions Implications: An Overview." *Global Environmental Change* 42 (January): 153–68. <https://doi.org/10.1016/j.gloenvcha.2016.05.009>.
- Steen-Olsen, Kjartan, Anne Owen, Edgar G. Hertwich, and Manfred Lenzen. 2014. "EFFECTS OF SECTOR AGGREGATION ON CO2 MULTIPLIERS IN MULTIREGIONAL INPUT–OUTPUT ANALYSES."

- Economic Systems Research* 26 (3): 284–302.
<https://doi.org/10.1080/09535314.2014.934325>.
- Veen, Anne van der. 2004. “Disasters and Economic Damage: Macro, Meso and Micro Approaches.” *Disaster Prevention and Management: An International Journal* 13 (4): 274–79.
<https://doi.org/10.1108/09653560410556483>.
- Ward, Hauke, Leonie Wenz, Jan C. Steckel, and Jan C. Minx. 2018. “Truncation Error Estimates in Process Life Cycle Assessment Using Input-Output Analysis: Truncation Error Estimates in Life Cycle Assessment.” *Journal of Industrial Ecology* 22 (5): 1080–91.
<https://doi.org/10.1111/jiec.12655>.
- Wiedmann, Thomas. 2009. “EDITORIAL: CARBON FOOTPRINT AND INPUT-OUTPUT ANALYSIS – AN INTRODUCTION.” *Economic Systems Research* 21 (3): 175–86.
<https://doi.org/10.1080/09535310903541256>.
- Wiedmann, Thomas, and Manfred Lenzen. 2018. “Environmental and Social Footprints of International Trade.” *Nature Geoscience* 11 (5): 314–21. <https://doi.org/10.1038/s41561-018-0113-9>.
- Wiedmann, Thomas O., Heinz Schandl, Manfred Lenzen, Daniel Moran, Sangwon Suh, James West, and Keiichiro Kanemoto. 2015. “The Material Footprint of Nations.” *Proceedings of the National Academy of Sciences* 112 (20): 6271–76. <https://doi.org/10.1073/pnas.1220362110>.
- Wiedmann, Thomas O., Heinz Schandl, and Daniel Moran. 2015. “The Footprint of Using Metals: New Metrics of Consumption and Productivity.” *Environmental Economics and Policy Studies* 17 (3): 369–88. <https://doi.org/10.1007/s10018-014-0085-y>.
- Wiedmann, Thomas, Harry C. Wilting, Manfred Lenzen, Stephan Lutter, and Viveka Palm. 2011. “Quo Vadis MRIO? Methodological, Data and Institutional Requirements for Multi-Region Input-Output Analysis.” *Ecological Economics* 70 (11): 1937–45.
<https://doi.org/10.1016/j.ecolecon.2011.06.014>.
- Willner, Sven Norman, Christian Otto, and Anders Levermann. 2018. “Global Economic Response to River Floods.” *Nature Climate Change* 8 (7): 594–98. <https://doi.org/10.1038/s41558-018-0173-2>.
- Wood, Richard, Troy R. Hawkins, Edgar G. Hertwich, and Arnold Tukker. 2014. “HARMONISING NATIONAL INPUT-OUTPUT TABLES FOR CONSUMPTION-BASED ACCOUNTING — EXPERIENCES FROM EXIOPOL.” *Economic Systems Research* 26 (4): 387–409.
<https://doi.org/10.1080/09535314.2014.960913>.
- World Bank. 2019. “GDP (Current US\$).” World Development Indicators.
<https://data.worldbank.org/indicator/NY.GDP.MKTP.CD>.
- Yang, Yi, Reinout Heijungs, and Miguel Brandão. 2017. “Hybrid Life Cycle Assessment (LCA) Does Not Necessarily Yield More Accurate Results than Process-Based LCA.” *Journal of Cleaner Production* 150 (May): 237–42. <https://doi.org/10.1016/j.jclepro.2017.03.006>.

Annexes

ANNEX I: MATCHING EORA ECONOMIC SECTORS WITH 2006 IPCC'S EMISSIONS CATEGORIES

This first annex describes the process of associating IPCC's energy categories to EORA 26's economic sectors. See §I.B.2.

1. Energy

The energy category covers a broad set of emissions sources related to energy extraction, transport, and use. IPCC defines three subcategories: Fuel Combustion Activities, Fugitive Emissions from Fuels and Carbon Dioxide Transport and Storage. The last category is currently not reported by countries. We therefore do not consider it.

a. Fuel Combustion Activities

The first subcategory is still large enough to concern most economic sectors. The first class of this subcategory refers to Energy Industries. It gathers Electricity and Heat Production (1A1a), Petroleum Refining (1A1b) and Manufacture of Solid Fuels and Other Energy Industries (1A1c). The first one relates to EORA's *Electricity, Gas and Water*, while we link the two formers to *Petroleum, Chemical and Non-Metallic Mineral Products*. The second class gathers emissions related to direct emissions from industrial sectors. We associate them with the corresponding economic sector in EORA. Note that we do not take in account the last subclass (1A2m, Non-specified industry). We assign the related emissions to the other industries. Transport (excluding international aviation) makes the third category and corresponds to EORA's *Transport*. IPCC names the forth category *Other Sectors*. It includes three subcategories: Commercial/Institutional, Residential and Agriculture/Forestry/Fishing/Fish Farms. This last subcategory divides once again in a few items that we assign conversely to *Agriculture* or *Fishing*. The second subcategory relates to EORA's *Final Consumption of Households*. The first category is challenging, because it represents decentralized emissions that we could assign to almost every economic sector. Nevertheless, this subclass belongs to the *fuel combustion* subcategory. Those emissions thus relate to the use of products from the *Petroleum, Chemical and Non-Metallic Mineral Products* sector. Since environmental applications of MRIO models mostly consist in assigning territorial emissions to the consumer that caused them, such approximation is acceptable. The same reasoning applies to the last class *Non-Specified* (1A5).

b. Fugitive Emissions from Fuels

The second subcategory of Energy regroups three subclasses, related to solid fuels (1B1), Oil and Natural Gas (1B2) and other Energy Production (1B3). *Mining & Quarrying* suits for the first and *Petroleum, Chemical and Non-Metallic Mineral Products* for the second. The last one regroups undefined emissions not considered in the previous classes from the energy sector. We assign those residual emissions to *Petroleum, Chemical and Non-Metallic Mineral Products*.

2. Industrial Processes and Product Use

The subcategories of the second category refer to numerous industries direct emissions. Most of them directly suit a unique EORA sector. 2F (*Product Uses as Substitutes for Ozone Depleting Substances*) embodies greenhouse gases used mostly in Air Conditioning devices. We therefore

allocate it to the *Electrical and Machinery* EORA sector. Subcategory 2G gathers *Other Product Manufacture and Use*. Even if IPCC defines numerous classes in this subcategory, available data do not allow distinguishing between those items. Still, most of those subclasses relate to the use of Electrical equipment. Therefore, we also bind it to EORA's *Electrical and Machinery*. The last subcategory (*Other*) gathers three classes: *Pulp and Paper Industry*, *Food and Beverages Industry* and *Other*. We omit the third class¹.

3. AFOLU

AFOLU is the acronym for *Agriculture, Forestry and Other Land Use*. However, PRIMAP does not account for land-use related changes. In practical, this category only contains emissions related to Agriculture and Livestock. Both belong to EORA's Agriculture economic sector.

4. Waste and Other

The classification of waste-related emissions is challenging, because all economic sectors emit waste, and EORA defines no category in charge of waste collection and treatment. Consequently, we omit this category. Emissions initially counted in PRIMAP's *Waste* category will consequently add to other sector's emissions, at the pro-rata of their proper account. The same logic applies to the fifth IPCC's category, which contains mostly indirect emissions of N₂O.

¹ 'Omitted' emissions are reallocated into the other categories of the same class, consistently with the computed repartition

ANNEX II: QUALITY ASSESSMENT OF THE EORA'S ENVIRONMENTAL INDICATORS

Following paragraphs investigate the consistence of several indicators introduced in EORA's environmental matrix. Due to the time-consuming dimension of those assessments, we only investigated the most relevant indicators. It relates to §II.B.2.

	TYPE	INDEX	SOURCE
POLLUANT EMISSIONS	Greenhouse Gases		
	- All - CO₂ - Others	1855 – 1857 ; 2606-2638 10 – 119 ; 1841 – 1854 ; 1858 ; 2520 – 2536 ; 2707 - 2715 120 – 1384 ; 2537 – 2604 ; 2639 – 2706	Interpolation ; PRIMAP ¹ Interpolation ; Unknown ; Unknown ; PRIMAP ¹ ; CDIAC – EDGAR – IEA Interpolation; Unknown; PRIMAP ¹ ;
PHYSICAL EXTRACTION	Local Polluants :		
	- Gases - Nitrogen and phosphorus emissions	1389 – 1828 2431 – 2436 ; 2470 – 2474	Interpolation SEDAC ; (Oita et al. 2016)
LAND USE	Water use	2475 – 2502	Water Footprint Network
	Primary Energy Use		
ECONOMY	- Total - By Energy source - HANPP²	1866 ; 2719 ; 2720 1 – 9 1385 - 1389	Interpolation : IEA Unknown Unknown
	Material Consumption	1865 – 1907	UNEP
LAND USE	Agricultural Production	1908 – 2423	FAOSTAT
	By culture type	2080 – 2251 ; 2424 - 2430	FAOSTAT
ECONOMY	By biotope	2439 – 2469	FAOSTAT
	Employment	1839 – 1840	ILO
ECONOMY	Value-added	1859 – 1865	Unknown
	GDP	2716 – 2717	IEA
ECONOMY	Population	2718	IEA

The large size of the database excludes analyzing the precise allocation of values between individual sectors all around the world. Instead, we compare values in EORA with their source, to ensure the consistency between data in EORA and reference database. Apart for a few exceptions, we therefore exclude data with unknown source. We also exclude agricultural data from the analysis, because they do not appear relevant for industrial applications.

1. Primary Energy Production

EORA provides the primary energy consumption of each sector, distinguishing nine energy sources:

¹ This indicator proceeds from a former incomplete integration of PRIMAP in EORA. Paragraph B.2.g describes the implementation of a corrected version of PRIMAP in EORA 26, at Index 2721-2729

² Human Appropriation of Net Primary Productivity ; Measure of the part of energy produced in nature captured by humankind. See <https://was-research.org/paper/an-introduction-to-human-appropriation-of-net-primary-productivity/>

- natural gas,
- coal,
- petroleum,
- nuclear electricity,
- hydroelectricity,
- geothermal electricity,
- wind electricity,
- solar, tide and wave electricity
- biomass and waste electricity.

Compared to IEA reference table¹, values coincide with 2010 data, with less than 10% differences for major sources of energy. Considering changes in electricity in recent years, data of this category appears a bit outdated, since there was no update of the table since around almost 10 years.

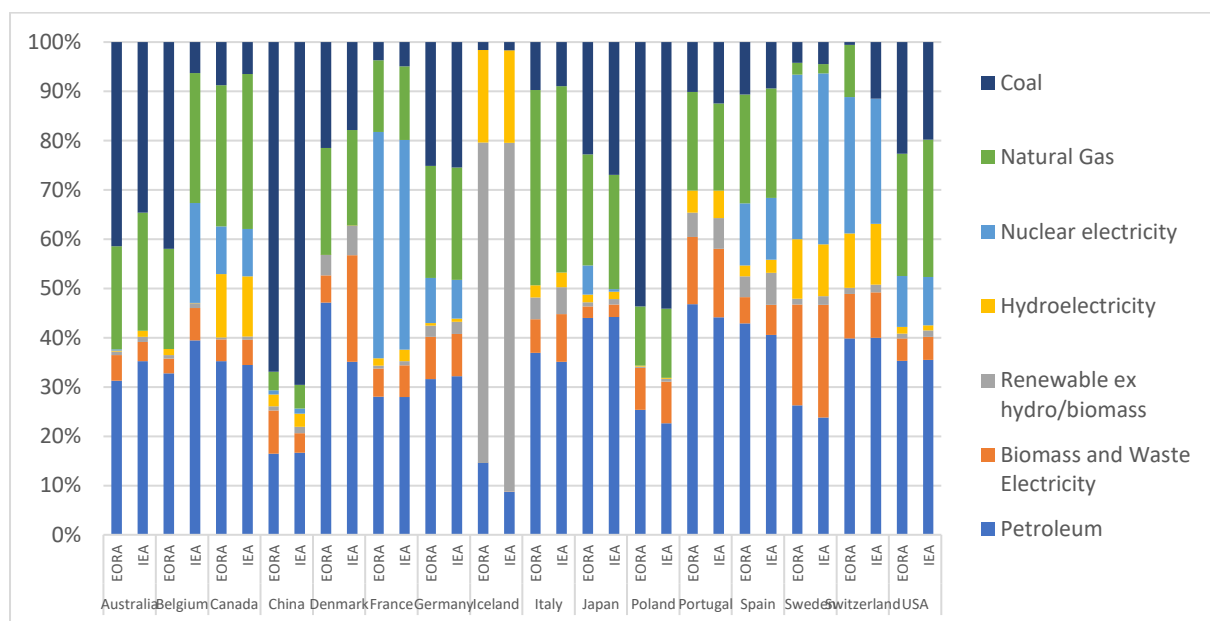


Figure 20 - Energy mix comparison of diverse countries between EORA 2013 and IEA 2016

While the mapping of some countries is unsatisfactory in EORA compared to reference data (especially in Belgium), available data still provide useful trend on the global energy mix of most countries. However, energy mix available in EORA 26 differs largely from Full EORA, and from IEA reference, underestimating by 65% biomass and waste Electricity supply, and by 22% solar supply compared to Full EORA. Concisely, this indicator provides a consistent energy mix for most countries in Full EORA, even if data last from 2010. It is however not consistent in EORA 26.

2. Employment

EORA data on employment proceed from the International Labor Organization (ILO). This indicator does not cover all countries: many African countries have no data available. Values are in thousands of Full Time Equivalents (FTE) employees and distinguish male from female employment. There is no open database for FTE statistics. We therefore compare available data in EORA with a

¹ <https://www.iea.org/statistics/>

custom measure, computed from OECD employment data¹ and a custom full time equivalent of 2080 hours per year.

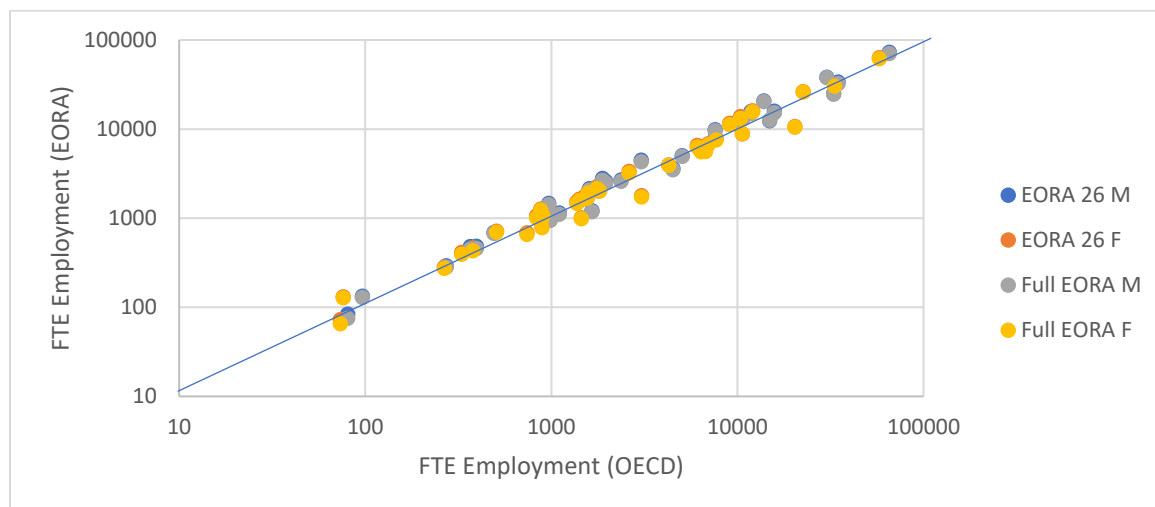


Figure 21 - Comparison of national Full Time Equivalent employment for year 2013

For each point of the graph, its coordinates correspond to the value estimated using OECD data and value observed in EORA. If both data were perfectly equal, all points would fit the line $Y = X$. It is almost the case here, both for male and female, and for both EORA models. Indicator looks then very consistent with externally reported data. Deviation from the axis $Y = X$ look small in account of the differences in methodology. Qualitatively, sectorial repartition also appears consistent, with agricultural sectors in China being the first employer of the world, followed by commercial business in India and China, then Chinese Construction, Public Administration and Education. Restricted to OECD members, our analysis underlines a good data quality for both models, even if this indicator omits many developing countries. In addition, there is no regular update of this data: values in EORA are constant between 2013 and 2015. However, employment market evolves in a relatively long-time span. Application of the model to employment related concern appears quite relevant.

3. Physical Extractions

EORA displays precise data on material extraction: there is 35 items, aggregated in four broad categories: biomass, metallic ores, non-metallic ores and fossil fuel. Data supported a publication in a peer-reviewed journal (T. O. Wiedmann et al. 2015), and primarily sources from the UNEP material flow dataset ("Global Material Flows Database" 2018), following the Economy-Wide Material Flow Accounting (EWMFA) principles (Fischer-Kowalski et al. 2011).

¹ <https://data.oecd.org/emp/hours-worked.htm#indicator-chart>

Annex II: Quality assessment of the EORA's environmental indicators

Category	Item
Biomass	Cereals
	Other crops
	Roots and tubers
	Sugar Crops
	Pulses
	Nuts
	Oil Bearing crops
	Vegetables
	Fruits
	Fibres
	Crop residues
	Grazed biomass
	Timber (industrial roundwood)
	Wood fuel and other extraction
Metallic Ores	Iron Ores
	Copper ores – gross ore
	Nickel ores – gross ore
	Lead ores – gross ore
	Zinc ores – gross ore
	Tin ores – gross ore
	Gold, silver, platinum and other precious metal ores – gross ore
	Bauxite and other aluminum ores – gross ore
	Uranium and thorium ores – gross ore
	Other metal ores – gross ore
Non-Metallic Ores	Ornamental or building stones
	Chalk and folomite
	Chemical and fertilizer minerals
	Salt
	Other mining quarrying products
	Non-metallic minerals – primarily construction
Fossil Fuel	Brown coal
	Hard coal
	Peat
	Crude oil and natural gas liquids
	Natural gas

Table 7 - List of material flow indicators

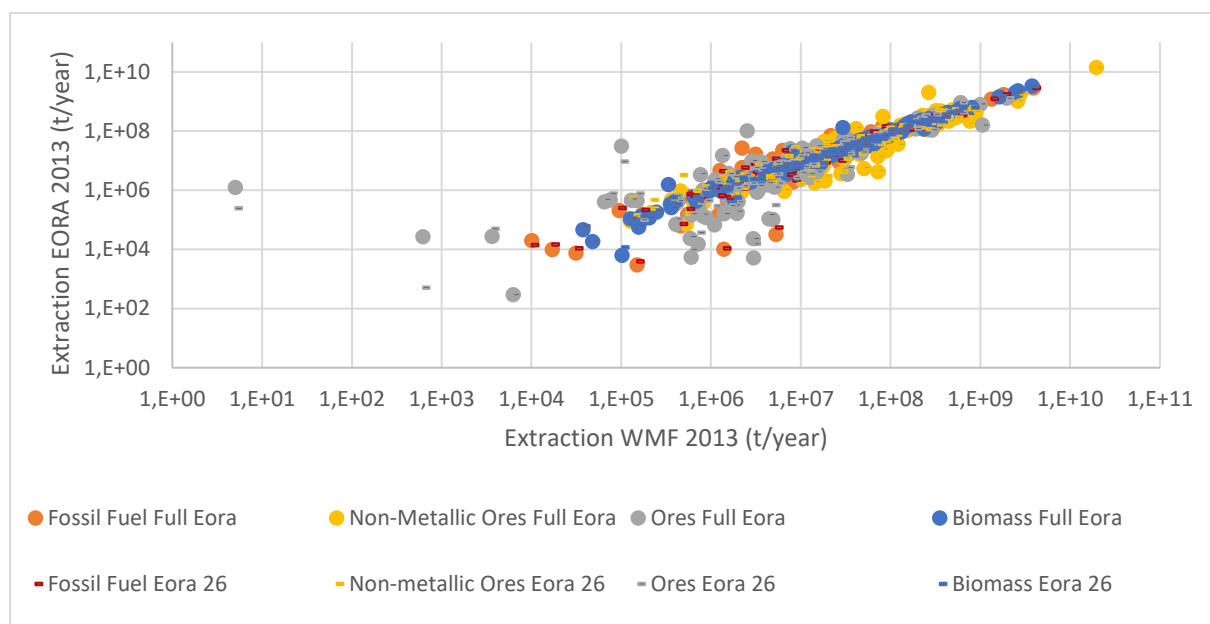


Figure 22 - Allocation of Material Extraction between countries in 2013

Allocation of physical extraction shows slight divergences between EORA and reference source. However, global trend is still close to axis $Y = X$. Deviation could process from the absence of updating of the database, or from methodological issues. Non-metallic and metallic ores display higher divergences than biomass and fossil fuels. This could come from a higher year-to-year variability. Allocation in EORA 26 is much comparable in quality to Full EORA. Quantitative difference for the total ranges from 10 to 20% for both models. Sectorial allocation appears satisfactory. Biomass extraction mostly relate to agriculture (76%) and food and beverages (19%) related sectors, while Mining and Quarrying sectors embody more than 98% of remaining extractions. Consequently, this group of indicators present slight deviation with external data sources in terms of territorial allocation. Considering the range of the differences, user should consider output using those indicators as order of magnitudes better than exact values.

4. Water Extraction

The Water Footprint Network¹ provides a database for water usage, distinguishing three categories. Blue water accounts for surface and groundwater, green water for precipitation and grey water for the quantity of water required to dilute pollutants to an acceptable level. EORA embodies WFN data from 2005, and supported a publication in a peer-reviewed journal (Lenzen et al. 2013).

¹ <https://waterfootprint.org/en/resources/waterstat/>

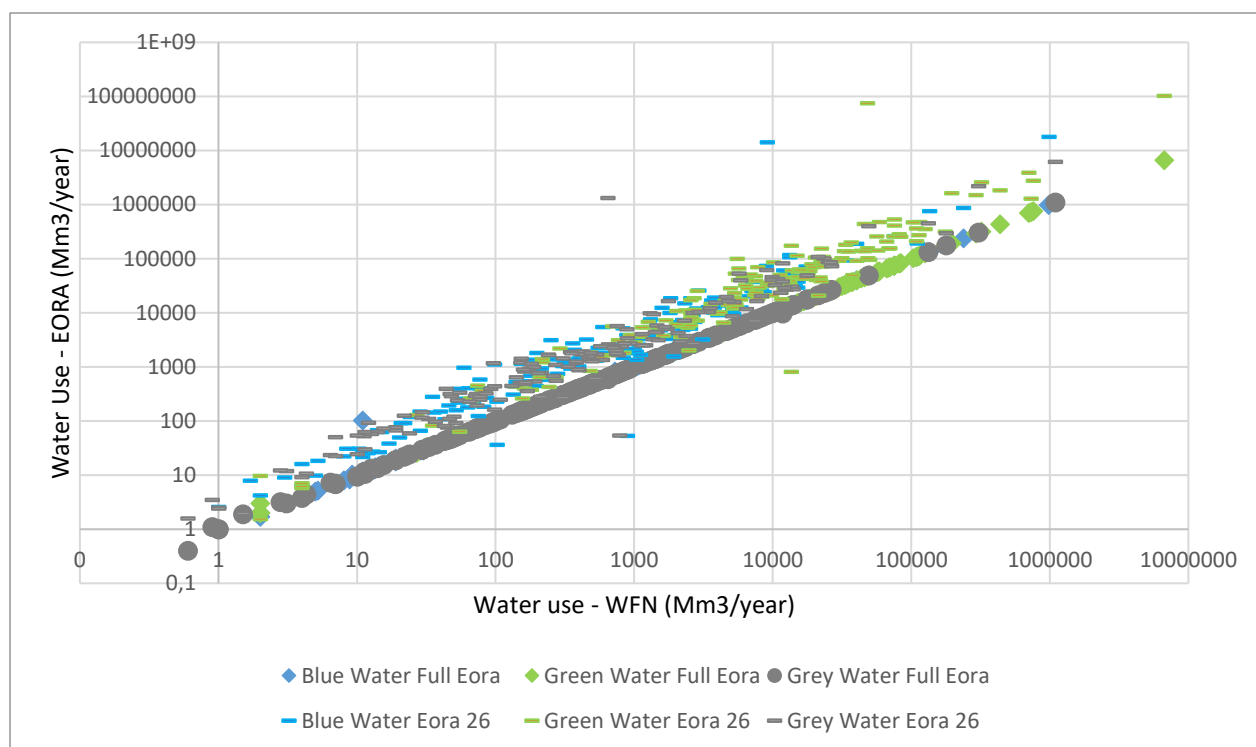


Figure 23 - Comparison of water use accounting

Data points for Full EORA almost perfectly fit the equation $Y = X$, no matter the type of extraction. Allocation in EORA 26 appears much less consistent. While all data relatively fits a straight line, this line lays half a decade over previous axis. This difference advocates for a strong systematic error in water use allocation in EORA 26. While the reduced model contains widely overestimated values, Full EORA database appears very close to reference data. While raw data advocates for strong confidence, the main drawback of this indicator rely in its lack of recent updating.

5. Nitrogen emissions

EORA gathers two sets of indicators for nitrogen emissions. One of them refers to agricultural uses, while the other one includes several nitrogen related components. We do not assess the first one, for lack of obvious application. The second one supported a recent academic publication (Oita et al. 2016). It gathers emissions of N₂O, NH₃, NO_x, and direct absorption by water bodies. There is no reference database for the last category of nitrogen pollution. We compare other emissions with Edgar¹, a main reference for greenhouse gases and air pollutant emissions.

¹ https://edgar.jrc.ec.europa.eu/overview.php?v=432_AP&SECURE=123

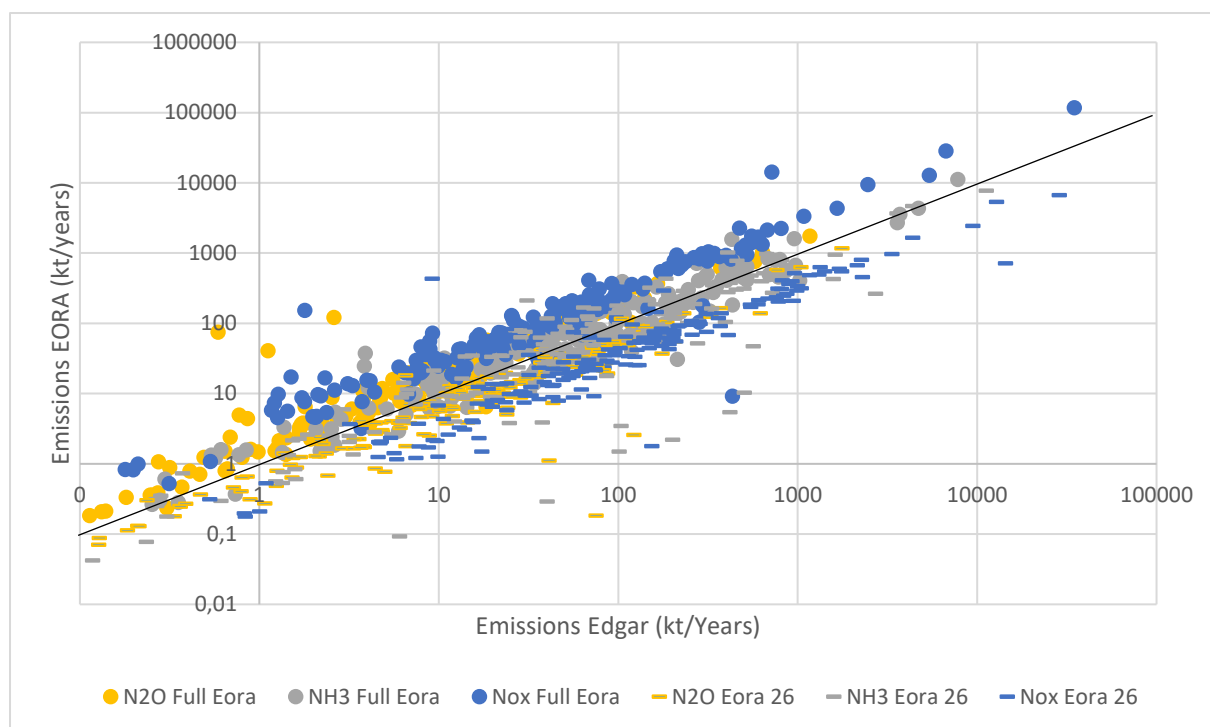


Figure 24 - Comparison of Nitrogen emissions

Comparison of EORA data to Edgar reference lights a medium adherence of the indicator to the source. There is a relatively good consistence in terms of ordination, but absolute values seem overestimated from half a decade in Full EORA and underestimated from half a decade in EORA 26. Some isolated points diverge from more than order of magnitude between both sources, including for some very important values. While divergences in small coefficients have often a limited influence on overall results, deviations in larger numbers may trigger important numerical errors in calculation. For this reason, this indicator comes with low certainty. Apart from numerical values, sectorial repartition appears relevant. Most emissions proceed from agricultural sector, with also noticeable contribution for NOx by transport and energy production sectors.

6. GHG emissions

EORA satellite matrix includes numerous GHG indicators, issued from different sources. Five main emissions set are available: PRIMAP, CDIAC, EDGAR, IEA, and an interpolated indicator computed from the different sources. CDIAC, EDGAR and IEA provide only CO2 emissions data, while PRIMAP and the interpolation provide various GHGs and air pollutants, with IPCC category classifications.

Annex II: Quality assessment of the EORA's environmental indicators

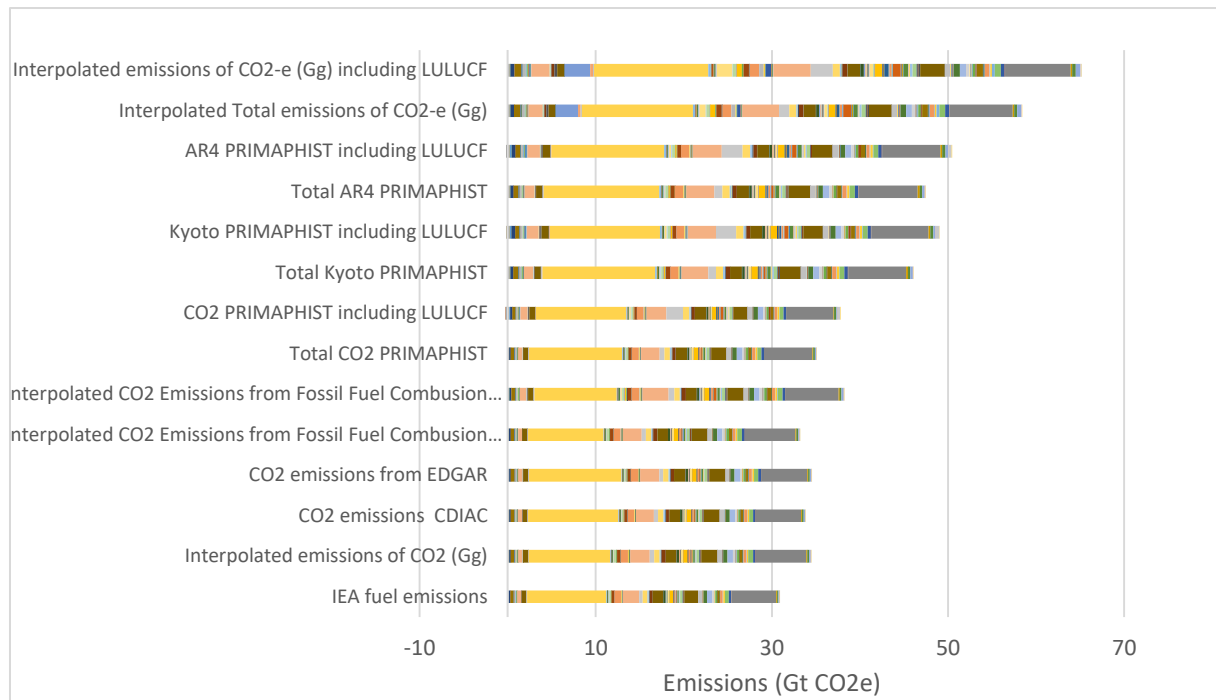


Figure 25 - Comparison of main GHG emission data in Full EORA 2013

While PRIMAP and the interpolations appear to be the most appealing source, they are in practical unusable, because of a poor sectorial repartition. Figure 6 illustrates the repartition of CO2 emissions among countries, following different indicators. Note that the first six indicators are aggregates of GHG emissions, while the others only consider CO2 emissions. Those indicates a volume of total emissions included between 45 and 65 Gt CO2 equivalent. Interpolated values nonetheless appear suspicious, with Central African Republic emitting as much emission as Russia. This obvious inconsistency impedes the use of this database. Conversely, sectorial repartition of PRIMAP emissions is also problematic, with no emissions associated with highly polluting energy sectors, such as the Chinese electric sector. IEA also provides surprising results for sectorial repartition, with rail transportation in the US being the second biggest CO2 emitter in the world. Therefore, only two data sources appear exploitable for CO2 related analysis, EDGAR and CDIAC. Both databases appear highly consistent one with another, with around 34 Gt CO2 for the whole world. We ultimately choose Edgar for the reference indicator, for two reasons. First, Edgar data support most carbon footprint applications of EORA (Moran and Wood 2014; K. Kanemoto et al. 2014; Jiborn et al. 2018). Secondly, 2015 data includes an update of Edgar, but not of CDIAC.

Section I.B.2) extensively describes the integration of fresh PRIMAP data into EORA and provides specific information on the quality of this mapping. It is globally highly consistent with the Full EORA mapping of EDGAR. We therefore consider it as a very good quality.

