

Software paper for submission to the Journal of Open Research Software
Please submit the completed paper to: editor.jors@ubiquitypress.com

(1) Overview

Title

Pymrio - a Python based Multi-Regional Input-Output Analysis toolbox

Paper Authors

1. Stadler, Konstantin

Paper Author Roles and Affiliations

Developer of Pymrio, Industrial Ecology Programme, Norwegian University of Science and Technology (NTNU)

Abstract

Pymrio is an open source tool for Environmentally Extended Multi-Regional Input-Output (EE MRIO) analysis developed in Python. It provides a high-level abstraction layer for global EE MRIO databases in order to simplify common EE MRIO data tasks. Among others, Pymrio includes parsers for several openly available EE MRIO databases (EXIOBASE v1 - v3, WIOD, Eora26, OECD-ICIO) as well as methods for production- and consumption-based accounts calculation, aggregation, stressor origin estimation and visualization. The use of a consistent storage format including meta data and modification history for MRIOs allows to exchange data with other analysis tools, aiming for an increased interoperability of Industrial Ecology analysis software.

Keywords

Sustainability Analysis, Multi Regional Input Output Analysis, Footprinting, Consumption-Based Accounting, EXIOBASE, WIOD, Eora26, OECD-ICIO, Python

Introduction

Environmentally Extended Multi-Regional Input-Output (EE MRIO) tables describe economic relationships within and between regions and their environmental repercussions.

They capture the full life-cycle impacts of goods and services across international supply chains and allow researchers to understand the environmental, social and economic consequences of consumption in today's globalised world. Research questions that can be answered using MRIO tables range from calculating the carbon footprint of nations [14, 19] and households [18], to analysing the impacts of international trade on biodiversity [10, 28, 37], to investigations of the material requirements of the current global economy [12]. In addition, EE MRIO analysis is increasingly used to inform policy makers about connections, trade-offs and synergies across environmental and social issues [13, 26, 36, 38].

The MRIO framework

The basic building blocks of EE MRIO models are (1) domestic input-output (IO) tables linked with (2) detailed bilateral trade data and extended with (3) environmental accounts [9].

The first building block, the domestic IO tables, depict the inter-industry relationship within an economy. These are typically square tables with column/row indices for all industries within a country. One column in these tables describes the inputs (products from other industries) required by one specific industry (or more general product production process) to produce its output. Arranging these column vectors into a two-dimensional matrix, the rows then depict all users of the output of a specific industry. Thus, the column of an IO table describes an industry in its role as user of products while the corresponding row describes its role as supplier.

In most cases these tables are further extended by sections representing the value added of each industry, taxes and subsidies as well as final consumption (private and government consumers as well as capital formation). Input-output tables following the United Nations System of National Accounts, UN SNA [6], system are inherently complete and consistent, thus capturing all economic activity within a country where the sum over all inputs into industries (including imports) equals the output of all industries (including exports). The data for domestic input output tables are typically gathered by industry surveys conducted by national statistical agencies as the tables form an integral part of the national accounts.

The pivotal point to derive global multi regional input output tables from these domestic ones is to connect them through detailed description of bilateral trade flows. These connect each industry/product sector of a specific country with all other sectors of all other countries, thus describing how the output of one sector (as export) is used as input (through imports) in any other country. Similar to domestic IO tables, global MRIOs are inherently complete and consistent, thus capturing all global trade flows and economic activity with two caveats: often multiple countries are aggregated to world regions and the tables usually only describe the official economy (thus reported economic activity) and exclude any shadow economic or financial flows due to criminal activity.

Since global MRIOs allow the tracking of products through the global supply chain they have been proven to be an ideal tool to also track the environmental burden associated with the production processes along the supply chains. In order to do so, one conceptually regards environmental effects like land use, emissions or material use as required inputs into the production process of a certain product. This links the environmental consequences due to the production of a product to the product itself. Fundamentally, these opens up two contrasting perspectives: Production-based environmental accounts, which describe the environmental consequences at the place of production and the consumption-based environmental accounts, which assigns the environmental consequences of production to the final consumer of the product. The later are also known as various kinds of footprints and became a stable tool to describe the environmental performance of countries [14, 17, 21].

The mathematical handling of these tables is described further below in the section "Mathematical Background".

Current status for MRIO analysis software

In contrast to other Industrial Ecology methodologies like Life Cycle Assessment (e.g. Brightway [23], openLCA [31]) or Material Flow Analysis (e.g. STAN [4]), few generally available analysis packages for (Multi-Regional) Input-Output tables are available [15]. One of the very few well documented and stand-alone packages for IO analysis, PyIO [2], is not available for Python 3 and does not provide parsers for current MRIO databases (the package has not been updated since 2011 but is still available at <http://www.real.illinois.edu/pyio>). Another framework, the MRIOLab suite [20, 22] takes a different approach: it provides a virtual lab for the compilation of MRIO tables and thereby streamlining the compilation of MRIO tables. The MRIOLab also includes functions for MRIO analysis. These, however, are not well suited for the analysis of MRIO tables compiled independent of the MRIOLab since the tables can not fully reproduced within the MRIOLab [24, 25]. As a consequence of the lack of a generic MRIO analysis toolkit, MRIO analysis today relies on often ad-hoc produced scripts and functions. This hinders reproducibility of results and the reuse of previous coding efforts. Here I present the open source tool Pymrio, a Python 3 package, which aims to close this method gap for EE MRIO analysis.

The article proceeds with a description of the architecture of Pymrio, including the mathematical background and implementation details. This is followed by a short tutorial with a simple use case for Pymrio. This tutorial is also available as a Jupyter notebook. The reuse potential and future development plans are pointed out at the end of the article.

Implementation and architecture

Mathematical Background

This section gives an overview about the mathematical background of EE MRIO analysis as used in Pymrio.

The focus of this section is on the specifics of the multi-regional aspects of input-output analysis. Note, however, that Pymrio can also be used for analysing simple single country input-output tables. Readers unfamiliar with input-output analysis are referred to the standard textbook for input-output analysis by Miller and Blair [5]. Exercises of this book are also used in the online documentation which also highlights how Pymrio can be used for teaching of input-output methods (https://pymrio.readthedocs.io/en/latest/notebooks/pymrio_directly_assign_attributes.html).

Generally, mathematical routines implemented in Pymrio follow the equations described below. If, however, a more computationally efficient mechanism was available this was preferred. In these cases, the original formula remains as comment in the source code. Mostly, this was the case when instead of a matrix multiplication with a diagonalized vector the operation can be efficiently executed using numpy broadcasting mechanisms [8].

The Input-Output analysis implemented in Pymrio follows the classic Leontief demand-style modeling [1]. To do so, MRIO tables describe the global inter-industries flows within and across countries for k countries with a transaction matrix

Z :

$$Z = \begin{pmatrix} Z_{1,1} & Z_{1,2} & \cdots & Z_{1,k} \\ Z_{2,1} & Z_{2,2} & \cdots & Z_{2,k} \\ \vdots & \vdots & \ddots & \vdots \\ Z_{k,1} & Z_{k,2} & \cdots & Z_{k,k} \end{pmatrix} \quad (1)$$

Each submatrix on the main diagonal ($Z_{i,i}$) represents the domestic interactions for each industry n . The off diagonal matrices ($Z_{i,j}$) describe the trade from region i to region j (with $i, j = 1, \dots, k$) for each industry. Accordingly, global final demand can be represented by

$$Y = \begin{pmatrix} Y_{1,1} & Y_{1,2} & \cdots & Y_{1,k} \\ Y_{2,1} & Y_{2,2} & \cdots & Y_{2,k} \\ \vdots & \vdots & \ddots & \vdots \\ Y_{k,1} & Y_{k,2} & \cdots & Y_{k,k} \end{pmatrix} \quad (2)$$

with final demand satisfied by domestic production in the main diagonal ($Y_{i,i}$) and direct import to final demand from country i to j by $Y_{i,j}$.

The global economy can thus be described by:

$$x = Ze + Ye \quad (3)$$

with e representing the summation vector (column vector with 1's of appropriate dimension) and x the gross output.

The direct requirement matrix A is given by multiplication of Z with the diagonalized and inverted gross output x :

$$A = Z\hat{x}^{-1} \quad (4)$$

Based on the linear economy assumption of the IO model, gross output x can then be determined for any arbitrary vector of final demand y by multiplying with the total requirement matrix (Leontief matrix) L .

$$x = (I - A)^{-1}y = Ly \quad (5)$$

IO systems can be extended with various extensions (satellite accounts). Among others these can represent factors of production (e.g. value added, employment) and environmental stressors associated with production. These direct factors, contained in matrix F , can be normalized to the output per sector x by

$$S = F\hat{x}^{-1} \quad (6)$$

Multipliers (total, direct and indirect, requirement factors for one unit of output) are then obtained by

$$M = SL \quad (7)$$

Total requirements (footprints in case of environmental requirements) for any given final demand vector y are given by

$$D_{cba} = My \quad (8)$$

Setting the domestically satisfied final demand $Y_{i,i}$ to zero ($Y_t = Y - Y_{i,j} \mid i = j$) allows to calculate the factors of production occurring abroad (embodied in imports)

$$D_{imp} = SLY_t \quad (9)$$

The factors of production occurring domestically to satisfy final demand in other countries is given by:

$$D_{exp} = S\widehat{LY}_te \quad (10)$$

where $\widehat{}$ indicates diagonalization of the resulting column-vector of the term underneath.

If parts of the environmental stressors occurring during the final use of product, these can be represented by F_Y (e.g. household emissions). These need to be added to the total production- and consumption-based accounts to obtain the total impacts per country. Total territorial requirements/emissions per region i are therefore given by summing over the stressors per sector ($0 \dots m$) plus the stressors directly linked to final consumption for all final demand categories ($0 \dots w$) of that region.

$$D_{pba}^i = \sum_{s=0}^m F_s^i + \sum_{c=0}^w F_Y^{i_c} \quad (11)$$

Similarly, total requirements (footprints in case of environmental requirements) per region i are given by summing the detailed footprint accounts and adding the aggregated final demand stressors.

$$D_{cba}^i = \sum_{s=0}^m D_{cba}^{i_s} + \sum_{c=0}^w F_Y^{i_c} \quad (12)$$

Internally, the summations are implemented with the *group-by* functionality provided by the Pandas package [7].

Aggregation

For the aggregation of the MRIO system the matrix B_k defines the aggregation matrix for regions and B_n the aggregation matrix for sectors.

$$B_k = \begin{pmatrix} b_{1,1} & b_{1,2} & \cdots & b_{1,k} \\ b_{2,1} & b_{2,2} & \cdots & b_{2,k} \\ \vdots & \vdots & \ddots & \vdots \\ b_{w,1} & b_{w,2} & \cdots & b_{w,k} \end{pmatrix} B_n = \begin{pmatrix} b_{1,1} & b_{1,2} & \cdots & b_{1,n} \\ b_{2,1} & b_{2,2} & \cdots & b_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ b_{x,1} & b_{x,2} & \cdots & b_{x,n} \end{pmatrix} \quad (13)$$

With w and x defining the aggregated number of countries and sectors, respectively. Entries b are set to 1 if the sector/country of the column belong to the aggregated

sector/region in the corresponding row and zero otherwise. The complete aggregation matrix B is given by the Kronecker product of B_k and B_n :

$$B = B_k \otimes B_n \quad (14)$$

This effectively arranges the sector aggregation matrix B_n as defined by the region aggregation matrix B_k . Thus, for each 0 entry in B_k a block $B_n * 0$ is inserted in B and each 1 corresponds to $B_n * 1$ in B .

The aggregated IO system can then be obtained by

$$Z_{agg} = BZB^T \quad (15)$$

and

$$Y_{agg} = BY(B_k \otimes I)^T \quad (16)$$

with I defined as the identity matrix with the size equal to the number of final demand categories per country.

Factors of production are aggregated by

$$F_{agg} = FB^T \quad (17)$$

and stressors occurring during final demand by

$$F_{Y_{agg}} = F_Y(B_k \otimes I)^T \quad (18)$$

Implementation

The main design principle of Pymrio is based on the concept that an EE MRIO system can be effectively represented as an object in an Object-Oriented Programming (OOP) language. In Pymrio, such an EE MRIO object consists of a core component describing the economic relationships grouped with a various number of components describing the environmental and/or social extensions (satellite accounts, see Figure 1). All components of the main object are in turn represented as objects, allowing to implement specific methods for each sub-component.

This architecture described above was implemented in Python 3.7. The various tables of the MRIO system are stored in Pandas DataFrames [7], therefore building upon a well-tested data-science framework. As a consequence, besides the specific methods implemented by Pymrio, the full functionality of Pandas and the underlying numpy framework [8] can be used to modify the MRIO data.

Methods implemented in Pymrio which go beyond basic Pandas functionality are accompanied by a corresponding test harness which ensures the formal correctness of the method. The full source code is hosted on a public code repository together with an extensive documentation and tutorials <https://github.com/konstantinstadler/pymrio>. Pymrio is openly available under the GNU General Public License v3.0.

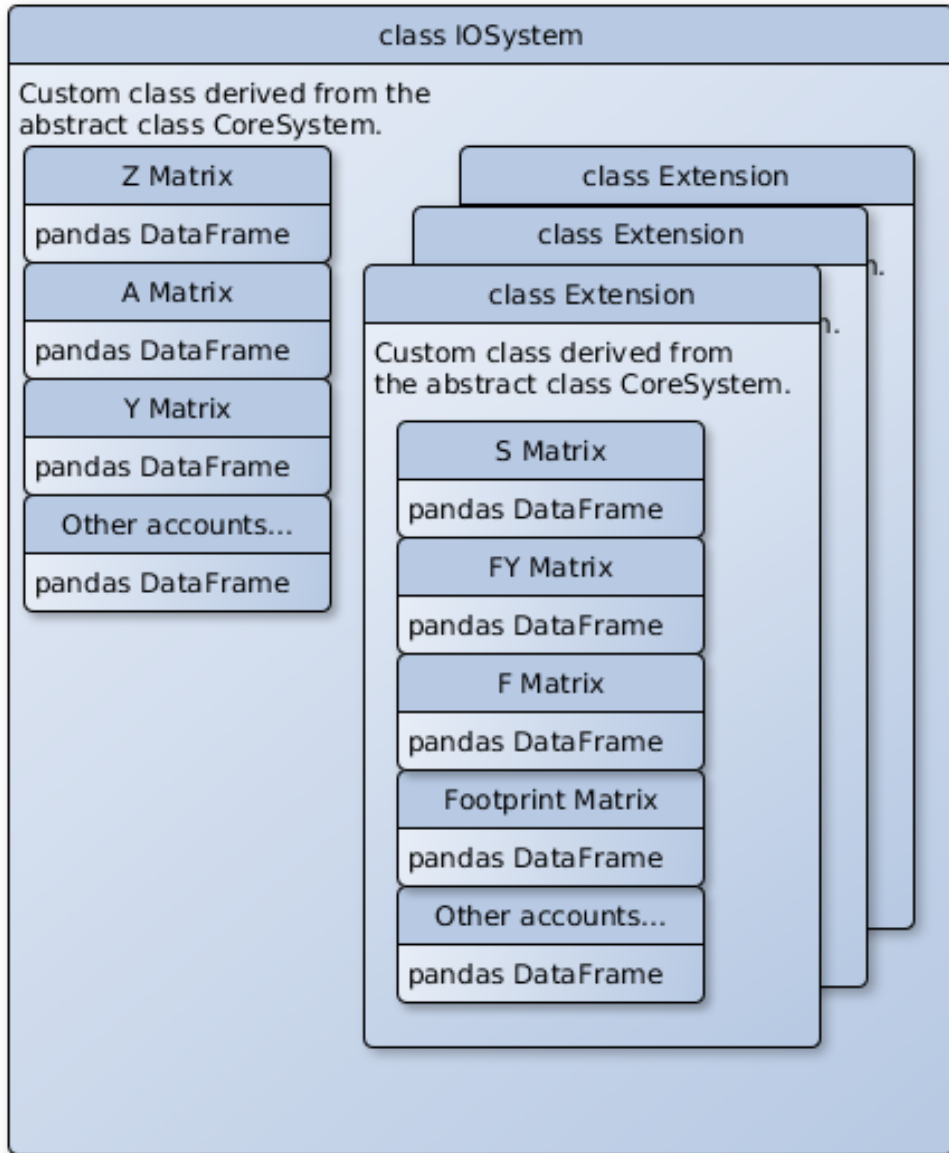


Figure 1: **Class diagram of the core Pymrio class.** The composite class IOSystem consists of the economic core with the actual data stored in Pandas DataFrames and a variable number of Extension classes. Each Extension consists of multiple Pandas DataFrames. Both, the IOSystem and Extension class are derived from an abstract CoreSystem class implementing the shared functionality of both classes. Class methods are not depicted here.

Parsing and Storage

To date, no standard way of storing MRIO databases has been defined. For example, the WIOD database [16] is provided as xlsx tables, whereas Eora [11] and EXIOBASE [35] use (compressed) csv tables. For the two latter, however, the approach differs as Eora26 uses pure numerical tables with separate files describing the headers, whereas EXIOBASE uses csv tables which include the headers. To

ease the use of different MRIO systems, Pymrio include parsers for the different formats. After parsing a MRIO system, Pymrio stores all data in a consistent way. For each component (core system and extension) data is stored in a separate folder. While the storage-format of the actual numerical data can be defined by the user, each storage folder also contains a json file (file_parameters.json) which contains information about the used format. Using the common json file format for storing the file meta data allows to easily and automatically import the data in other programming environments. By default, each table is stored as a tab-separated text file format, including row and column headers. The folder with the economic core also contains a file named “metadata.json” which includes information about version, name and system (industry-by-industry or product-by-product) as well as a record of modifications to the particular MRIO system (including when it was downloaded, applied aggregations, removal/addition of extensions, etc.).

Usage

The following section provides a quick start guide for using Pymrio beginning at the installation followed by a basic input-output calculation example. For the example here, the WIOD MRIO database [16] is used. However, after downloading and parsing the database the same methods are available for any EE MRIO system.

The code examples here are included as interactive Jupyter notebook tutorial in the SI <https://git.io/fjjUk>. As the code example here only show the code inputs, refer to the notebook to see the output of a specific command. An cloud-based virtual environment with the code example can be run in the BinderHub available at <https://cutt.ly/vwm0pA3>.

Pymrio is a Python [33] package, Python version ≥ 3.7 is required. The Pymrio package is hosted on PyPI [32] and the Anaconda Cloud [30]. Therefore, you can either use

```
pip install pymrio --upgrade
```

or

```
conda install -c conda-forge pymrio
```

to install Pymrio and all required packages.

Pymrio can then be used in any Python programming environment.

Throughout the code examples below, it is assumed that Pymrio is imported as follows:

```
import pymrio
```

First, the Pymrio MRIO download function is used to get the WIOD MRIO database with:

```
raw_wiod_path = '/tmp/wiod/raw'
pymrio.download_wiod2013(storage_folder=raw_wiod_path,
                        years=[2008])
```

This downloads the 2008 MRIO table from WIOD. Omitting the year parameter would result in a download of all years. The function returns a Pymrio meta

data object, which gives information about the WIOD version, system (in this case industry-by-industry) and records about from where the data was received (see SI cell 6).

To parse the database into a Pymrio object use:

```
wiod = pymrio.parse_wiod(raw_wiod_path, year=2008)
```

The available data can be explored by for example

```
wiod.get_sectors()
wiod.get_regions()
```

The transaction matrix can be inspected with

```
wiod.Z
```

which returns a panda DataFrame with the recorded monetary flows.

WIOD includes several extensions, which are stored as sub-objects (see Figure 1) in Pymrio. For example, in order to see the AIR emissions provided by WIOD:

```
wiod.AIR.F
```

WIOD, however, does neither provide any normalized data (A-matrix, satellite account coefficient data) nor any consumption-based accounts (footprints).

In order to calculate them, one could go through all the missing data and compute each account. Pymrio provides the required functions, for example to calculate the A-matrix:

```
x = pymrio.calc_x(Z=wiod.Z, Y=wiod.Y)
A = pymrio.calc_A(Z=wiod.Z, x=x)
```

Alternatively, Pymrio provides a function which identifies all missing accounts and calculates them:

```
wiod.calc_all()
```

At this point, a basic EE MRIO analysis is accomplished. For example, the regional consumption-based accounts of the AIR emissions are now given by:

```
wiod.AIR.D_cba_reg
```

Units are stored separately in

```
wiod.AIR.unit
```

Pymrio can be linked with the country converter coco [27] to ease the aggregation of MRIO and results into different classifications. Using the country converter, WIOD can easily be aggregated into EU and non-EU countries with singling out Germany and the UK by:

```
import country_converter as coco
wiod.aggregate(region_agg = coco.agg_conc(
    original_countries='WIOD',
    aggregates=[{ 'DEU': 'DEU', 'GBR': 'GBR' }, 'EU'],
    missing_countries='Other',
    merge_multiple_string=None))
wiod.rename_regions({ 'EU': 'Rest_of_EU' }
```

To visualize the results for example for CH₄ the matplotlib framework [3] can be used (Figure 2):

```
import matplotlib.pyplot as plt
with plt.style.context('ggplot'):
    wiod.AIR.plot_account('CH4')
    plt.savefig('airch4.png', dpi=300)
    plt.show()
```

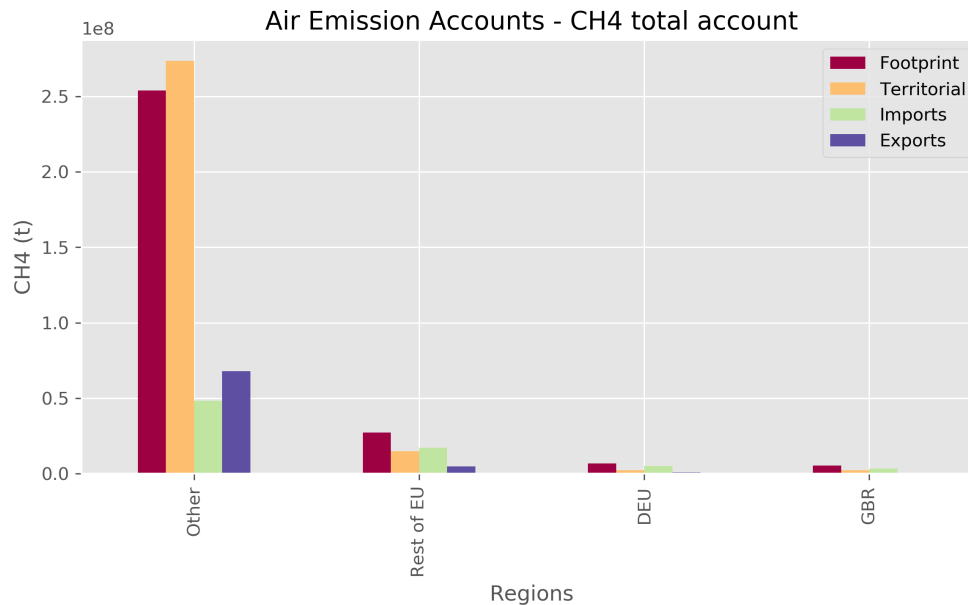


Figure 2: **CH₄ emissions of Germany (DEU), the UK (GBR), Rest of the EU and Other countries** This figure was produced with Pymrio and matplotlib after aggregating the WIOD countries into the three regions specified above

To calculate the source (in terms of regions and sectors) of a certain stressor or impact driven by consumption, one needs to diagonalize this stressor/impact.

This can be done with Pymrio by:

```
diag_CH4 = wiod.AIR.diag_stressor('CH4')
```

and be reassigned to the aggregated WIOD system:

```
wiod.CH4_source = diag_CH4
```

In the next step the automatic calculation routine of Pymrio is called again to compute the missing accounts in this new extension:

```
wiod.calc_all()
```

The diagonalized CH₄ data now shows the source and destination of the specified stressor (CH₄):

```
wiod.CH4_source.D_cba
```

In this square consumption-based accounts matrix, every column represents the amount of stressors occurring in each region - sector driven by the consumption stated in the column header. Conversely, each row states where the stressor impacts occurring in the row are distributed to (from where they are driven).

If only one specific aspect of the source is of interest for the analysis, the footprint matrix can easily be aggregated with the standard Pandas *groupby* function. For example, to aggregate to the source and receiving region of the stressor:

```
CH4_source_reg = wiod.CH4_source.D_cba.groupby(
    level='region', axis=0).sum().groupby(
    level='region', axis=1).sum()
```

Which can then be visualised using the seaborn heatmap [29] with (Figure 3):

```
import seaborn as sns
CH4_source_reg.columns.name = 'Receiving_region'
CH4_source_reg.index.name = 'Source_region'
sns.heatmap(CH4_source_reg, vmax=5E6,
            annot=True, cmap='YlOrRd', linewidths=0.1,
            cbar_kws={'label': 'CH4 emissions (t)',
                    'wiod.CH4_source.unit.unit[0]})})
plt.show()
```

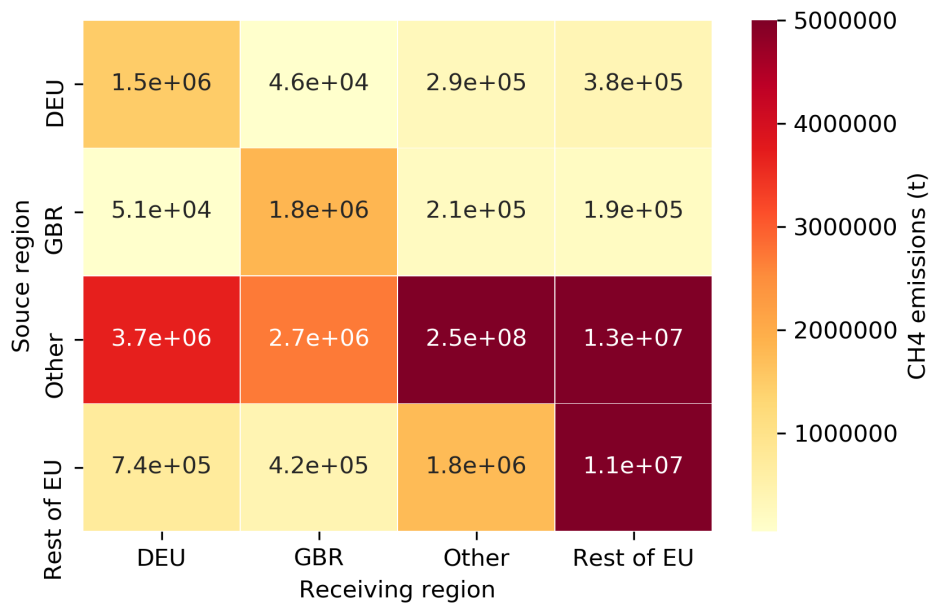


Figure 3: **CH₄ emissions source and destination** A substantial share of CH₄ originating in the Rest of the World region are exported into Germany and the UK. This figure was produced with Pymrio and seaborn.

Storing the MRIO database can be done with

```
storage_path = '/tmp/wiod/aly'
wiod.save_all(storage_path)
```

from where it can be received subsequently by:

```
wiod = pymrio.load_all(storage_path)
```

The meta attribute of Pymrio mentioned at the beginning kept track of all modifications of the system. This can be shown with:

```
wiod.meta
```

Custom notes can be added to the history with:

```
wiod.meta.note("Custom_note")
```

The history of the meta data can be filtered for specific entries like:

```
wiod.meta.file_io_history
```

This tutorial gave a short overview about the basic functionality of Pymrio. For more information about the capabilities of Pymrio check the online documentation at <http://pymrio.readthedocs.io> [34]

Quality control

All basic mathematical functions of Pymrio, as described in the section Mathematical Background above, are unit-tested against published results extracted from a classic input-output textbook [5]. Additional unit tests validate various components of the aggregation, file IO and other utility methods. Beyond the unit tests, Pymrio includes a small pseudo MRIO system which is used for a full regression test of the package. Currently, the test coverage is at around 90 %.

All tests are implemented in pytest and users can verify the correctness of Pymrio, after installing pytest, by

```
py.test -v
```

in the root of the local copy of Pymrio.

In addition, Pymrio uses the continuous integration platform Travis CI for automatic testing after each change of the code base uploaded to the source repository. After each build, test coverage is automatically calculated using the coveralls platform <https://coveralls.io>.

The Pymrio source code follows the PEP 8 specifications as implemented by the Black code formatter (<https://github.com/psf/black>). Accordance with it is tested through the continuous integration for each code contribution. More information on the code style and required tests for acceptance are outlined in the CONTRIBUTING file in the root of code repository.

The Pymrio documentation is build using the Sphinx Python Documentation Generator and hosted on readthedocs (<http://pymrio.readthedocs.io>).

After each change to the master branch, the API references in the documentation are automatically updated based on the description provided in the numpy style docstrings, thus keeping the documentation and code base in sync. Furthermore, the code examples and tutorials given in the documentation are implemented as Jupyter notebooks and are recalculated for each release, thus also serving as regression tests for the documented Pymrio functionality.

Updated results of the Travis CI tests as well as the Sphinx documentation rebuild are indicated at the beginning of the readme file at the source repository.

Contributors to the Pymrio code based are advised to adhere to the testing and code standards established for Pymrio. Further details can be found in the Contribution section of the online documentation.

(2) Availability

Operating system

GNU/Linux, Mac OSX, Windows and any other operating systems running Python with the SciPy stack.

Programming language

Pymrio was built in Python 3 and currently (Pymrio version 0.4.2) tested for Python 3.7, 3.8 and 3.9.

Additional system requirements

Pymrio runs on every system capable of running the Python SciPy stack. The actual memory requirements depend on the MRIO database to be analysed with Pymrio. For example, for EXIOBASE [35] a minimum of 8 GB RAM are required.

Dependencies

For the current Pymrio version 0.4.2:

- pandas \geq 1.0.0
- numpy \geq 1.13.4
- matplotlib \geq 2.0.0
- requests \geq 2.18
- xlrd \geq 1.1.0
- xlrd \geq 3.0.0
- docutils \geq 0.14

The main dependency of Pymrio is Pandas and future versions of Pymrio will follow the developmental changes in Pandas.

The file requirements.txt in the source repository contains an up-to-date list of all requirements.

For development and unit testing the pytest module and black code formatter are required. The requirements_test.txt file and alternatively the environment_dev.yml file (for using conda environments) outline all dependencies necessary for testing and development.

Software location:**Archive****Name:** Zenodo**Persistent identifier:** <https://doi.org/10.5281/zenodo.1146054>**Licence:** GPL v3**Publisher:** Konstantin Stadler**Version published:** 0.4.2 and earlier versions. The DOI above always resolves to the latest version, previous versions can be identified with separate DOIs (see versions sections on the Zenodo repository page).**Date published:** 19/11/20 (version 0.4.2)**Code repository****Name:** Github (Pymrio is also hosted on pypi and anaconda cloud)**Persistent identifier:** <https://github.com/konstantinstadler/pymrio>**Licence:** GPL v3**Date published:** 19/11/20 (version 0.4.2)**Emulation environment****Name:** MyBinder Jupyter Notebook of the tutorial included above**Persistent identifier:** https://mybinder.org/v2/gh/konstantinstadler/pymrio_article/master?filepath=%2Fnotebook%2Fpymrio-tutorial-for-wiod.ipynb**Licence:** CC BY 4.0**Date published:** 18/01/18**Language**

English

(3) Reuse potential

Pymrio contains functionality aimed at professional MRIO analysts and sustainability scientists, but might be useful to anyone doing environmental and/or economic analysis. As such, Pymrio is one key component in the Industrial Ecology analysis software framework [15]. With the other components it shares the ambition to improve usability, interoperability, and collaboration between Industrial Ecology and sustainability research Python packages. The main motivation for starting the project was to build a common interface for handling different MRIO databases, but through the years the scope extended to include visualization, reporting and data provenance tracking capabilities. Future development plans include further visualization possibilities, parser for additional MRIO models and extended analysis capabilities like structural decomposition and structural path analysis. Being an open source project, this includes an invitation to fellow researchers to join these coding efforts.

The primary communication channel for Pymrio is the GitHub source repository, in particular the Issue Tracker there. I strongly encouraged to not only use the Issue Tracker for bug reporting but for all questions, comments, and suggestions regarding the project. Pymrio follows an “issue driven development” style. This means that the first step for any modifications or enhancements to Pymrio are to file an issue describing the planned changes. This allows us to discuss changes

before the actual programming and gives us the chance to identify synergies across ongoing efforts and avoid potential double work. Finished modification should then be submitted as pull request. Further information about open points and code style can be found in the contributing.rst file at the source repository.

Acknowledgements

Special thanks to Guillaume Majeau-Bettez (Polytechnique Montréal / Norwegian University of Science and Technology), Radek Lonka (Norwegian University of Science and Technology), Richard Wood (Norwegian University of Science and Technology) and Stefan Pauliuk (University of Freiburg) for discussing ideas and the scope of Pymrio as well as implementation details. I am much obliged also to all people already using Pymrio and providing feedback and improvement suggestions.

Funding statement

The development of Pymrio was funded by the European Commission under the DESIRE Project (grant no.: 308552) as well as by The Research Council of Norway IKTPLUSS project HiTEA (grant no.: 302830).

Competing interests

The authors declare that they have no competing interests.

References

1. Leontief, W. Environmental Repercussions and the Economic Structure: An Input-Output Approach. *The Review of Economics and Statistics* **52**, 262–271. ISSN: 0034-6535. JSTOR: 1926294 (Aug. 1, 1970).
2. Nazara, S., Guo, D., Hewings, G. J. & Dridi, C. *PyIO - Input-Output Analysis with Python* 03-T-23 (The University of Illinois at Urbana-Champaign, Illinois, USA, 2003). <http://www.real.illinois.edu/pyio/>.
3. Hunter, J. D. Matplotlib: A 2D Graphics Environment. *Computing in Science & Engineering* **9**, 90–95. ISSN: 1521-9615 (2007).
4. Cencic, O. & Rechberger, H. Material Flow Analysis with Software STAN. *Journal of Environmental Engineering and Management* **18**, 3 (2008).
5. Miller, R. E. & Blair, P. D. *Input-Output Analysis: Foundations and Extensions* ISBN: 978-0-521-51713-3 0-521-51713-3 0-521-73902-0 978-0-521-73902-3 (Cambridge University Press, Cambridge [England]; New York, 2009).
6. United Nations, European Commission, International Monetary Fund, Organisation for Economic Co-operation and Development & World Bank. *System of National Accounts 2008* ISBN: 978-92-1-161522-7 92-1-161522-4 (United Nations, New York, 2009).
7. McKinney, W. *Data Structures for Statistical Computing in Python* in. Proceedings of the 9th Python in Science Conference (2010), 51–56. <http://conference.scipy.org/proceedings/scipy2010/mckinney.html> (2018).
8. Van der Walt, S., Colbert, S. C. & Varoquaux, G. The NumPy Array: A Structure for Efficient Numerical Computation. *Computing in Science & Engineering* **13**, 22–30. ISSN: 1521-9615 (Mar. 2011).

9. Wiedmann, T., Wilting, H. C., Lenzen, M., Lutter, S. & Palm, V. Quo Vadis MRIO? Methodological, Data and Institutional Requirements for Multi-Region Input–Output Analysis. *Ecological Economics* **70**, 1937–1945. ISSN: 0921-8009 (Sept. 15, 2011).
10. Lenzen, M. *et al.* International Trade Drives Biodiversity Threats in Developing Nations. *Nature* **486**, 109–112. ISSN: 0028-0836, 1476-4687 (June 6, 2012).
11. Lenzen, M., Moran, D., Kanemoto, K. & Geschke, A. Building Eora: A Global Multi-Region Input–Output Database at High Country and Sector Resolution. *Economic Systems Research* **25**, 20–49. ISSN: 0953-5314 (2013).
12. Wiedmann, T. O. *et al.* The Material Footprint of Nations. *Proceedings of the National Academy of Sciences*, 201220362. ISSN: 0027-8424, 1091-6490. pmid: 24003158 (Sept. 3, 2013).
13. Simas, M. S., Golsteijn, L., Huijbregts, M. A. J., Wood, R. & Hertwich, E. G. The “Bad Labor” Footprint: Quantifying the Social Impacts of Globalization. *Sustainability* **6**, 7514–7540 (Oct. 24, 2014).
14. Tukker, A. *et al.* *The Global Resource Footprint of Nations. Carbon, Water, Land and Materials Embodied in Trade and Final Consumption Calculated with EXIOBASE 2.1*. 1000 vols. ISBN: 978-3-200-03637-6. <http://exiobase.eu/9-home/27-creea-booklet> (Leiden/Delft/Vienna/Trondheim, 2014).
15. Pauliuk, S., Majeau-Bettez, G., Mutel, C. L., Steubing, B. & Stadler, K. Lifting Industrial Ecology Modeling to a New Level of Quality and Transparency: A Call for More Transparent Publications and a Collaborative Open Source Software Framework. *Journal of Industrial Ecology* **19**, 937–949. ISSN: 1530-9290 (Dec. 1, 2015).
16. Timmer, M. P., Dietzenbacher, E., Los, B., Stehrer, R. & de Vries, G. J. An Illustrated User Guide to the World Input–Output Database: The Case of Global Automotive Production. *Review of International Economics*, n/a–n/a. ISSN: 1467-9396 (May 1, 2015).
17. Giljum, S. *et al.* Identifying Priority Areas for European Resource Policies: A MRIO-Based Material Footprint Assessment. *Journal of Economic Structures* **5**, 1. ISSN: 2193-2409 (June 6, 2016).
18. Ivanova, D. *et al.* Environmental Impact Assessment of Household Consumption. *Journal of Industrial Ecology* **20**, 526–536. ISSN: 1530-9290 (June 1, 2016).
19. Kanemoto, K., Moran, D. & Hertwich, E. G. Mapping the Carbon Footprint of Nations. *Environmental Science & Technology* **50**, 10512–10517. ISSN: 0013-936X, 1520-5851 (Oct. 4, 2016).
20. Geschke, A. & Hadjikakou, M. Virtual Laboratories and MRIO Analysis – an Introduction. *Economic Systems Research* **29**, 143–157. ISSN: 0953-5314 (Apr. 3, 2017).
21. Ivanova, D. *et al.* Mapping the Carbon Footprint of EU Regions. *Environmental Research Letters* **12**, 054013. ISSN: 1748-9326 (2017).

22. Lenzen, M. *et al.* The Global MRIO Lab – Charting the World Economy. *Economic Systems Research* **29**, 158–186. ISSN: 0953-5314 (Apr. 3, 2017).
23. Mutel, C. Brightway: An Open Source Framework for Life Cycle Assessment. *The Journal of Open Source Software* **2**, 236. ISSN: 2475-9066 (Apr. 19, 2017).
24. Rahman, M. D. A. *et al.* A Flexible Adaptation of the WIOD Database in a Virtual Laboratory. *Economic Systems Research* **29**, 187–208. ISSN: 0953-5314 (Apr. 3, 2017).
25. Reyes, R. C. *et al.* The Virtual IELab – an Exercise in Replicating Part of the EXIOBASE V.2 Production Pipeline in a Virtual Laboratory. *Economic Systems Research* **29**, 209–233. ISSN: 0953-5314 (Apr. 3, 2017).
26. Simas, M., Pauliuk, S., Wood, R., Hertwich, E. G. & Stadler, K. Correlation between Production and Consumption-Based Environmental Indicators: The Link to Affluence and the Effect on Ranking Environmental Performance of Countries. *Ecological Indicators* **76**, 317–323. ISSN: 1470-160X (May 2017).
27. Stadler, K. The Country Converter Coco - a Python Package for Converting Country Names between Different Classification Schemes. *The Journal of Open Source Software* **2**. ISSN: 2475-9066 (Aug. 3, 2017).
28. Veronesi, F., Moran, D., Stadler, K., Kanemoto, K. & Wood, R. Resource Footprints and Their Ecosystem Consequences. *Scientific Reports* **7**, 40743. ISSN: 2045-2322 (Jan. 23, 2017).
29. Waskom, M. *et al.* Mwaskom/Seaborn: V0.8.1 (September 2017). *Zenodo* (Sept. 3, 2017).
30. Inc., A. *Anaconda Cloud* <https://anaconda.org/about> (2018).
31. openLCA. *openLCA Modeling Suite* — *openLCA.Org* <http://www.openlca.org/openlca/> (2018).
32. PyPI. *PyPI - the Python Package Index : Python Package Index* <https://pypi.python.org/pypi> (2018).
33. Python. *Official Python Webpage* <https://www.python.org/> (2018).
34. Stadler, K. *Pymrio - Multi Regional Input Output Analysis in Python* — *Pymrio 0.3.3 Documentation* <http://pymrio.readthedocs.io/en/latest/#> (2018).
35. Stadler, K. *et al.* EXIOBASE 3: Developing a Time Series of Detailed Environmentally Extended Multi-Regional Input-Output Tables. *Journal of Industrial Ecology* **22**, 502–515. ISSN: 1530-9290 (2018).
36. Vivanco, D. F., Wang, R. & Hertwich, E. Nexus Strength: A Novel Metric for Assessing the Global Resource Nexus. *Journal of Industrial Ecology* **22**, 1473–1486. ISSN: 1530-9290 (Dec. 1, 2018).
37. Marques, A. *et al.* Increasing Impacts of Land Use on Biodiversity and Carbon Sequestration Driven by Population and Economic Growth. *Nature Ecology & Evolution*, 1. ISSN: 2397-334X (Mar. 4, 2019).

38. Wood, R. *et al.* Beyond Peak Emission Transfers: Historical Impacts of Globalization and Future Impacts of Climate Policies on International Emission Transfers. *Climate Policy* **20**, S14–S27. ISSN: 1469-3062 (sup1 Apr. 8, 2020).

Copyright Notice

Authors who publish with this journal agree to the following terms:

Authors retain copyright and grant the journal right of first publication with the work simultaneously licensed under a Creative Commons Attribution License that allows others to share the work with an acknowledgement of the work's authorship and initial publication in this journal.

Authors are able to enter into separate, additional contractual arrangements for the non-exclusive distribution of the journal's published version of the work (e.g., post it to an institutional repository or publish it in a book), with an acknowledgement of its initial publication in this journal.

By submitting this paper you agree to the terms of this Copyright Notice, which will apply to this submission if and when it is published by this journal.