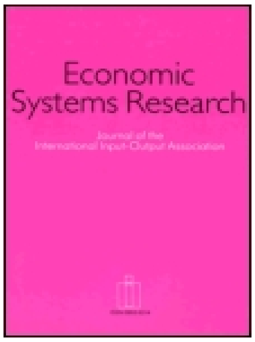


Economic Systems Research

ISSN: 0953-5314 (Print) 1469-5758 (Online) Journal homepage: https://www.tandfonline.com/loi/cesr20

CONVERGENCE BETWEEN THE EORA, WIOD, EXIOBASE, AND OPENEU'S CONSUMPTION-BASED CARBON ACCOUNTS

Daniel Moran & Richard Wood

Economic Systems Research, 2014
Vol. 26, No. 3, 245–261, http://dx.doi.org/10.1080/09535314.2014.935298

CONVERGENCE BETWEEN THE EORA, WIOD, EXIOBASE, AND OPENEU'S CONSUMPTION-BASED CARBON ACCOUNTS

DANIEL MORAN* and RICHARD WOOD

Industrial Ecology Programme and Department of Energy and Process Engineering, Norwegian University of Science and Technology (NTNU), Trondheim, Norway

(*Received 27 November 2013; accepted 12 June 2014*)

In this paper, we take an overview of several of the biggest independently constructed global multi-regional input–output (MRIO) databases and ask how reliable and consonant these databases are. The key question is whether MRIO accounts are robust enough for setting environmental policies. This paper compares the results of four global MRIOs: Eora, WIOD, EXIOBASE, and the GTAP-based OpenEU databases, and investigates how much each diverges from the multi-model mean. We also use Monte Carlo analysis to conduct sensitivity analysis of the robustness of each accounts' results and we test to see how much variation in the environmental satellite account, rather than the economic structure itself, causes divergence in results. After harmonising the satellite account, we found that carbon footprint results for most major economies disagree by $<10\%$ between MRIOs. Confidence estimates are necessary if MRIO methods and consumption-based accounting are to be used in environmental policy-making at the national level.

Keywords: MRIO; Footprint; CBA; Monte Carlo; Uncertainty; Reliability; Confidence

1. INTRODUCTION

Consumption-based accounts (CBA) built using global multi-regional input–output (MRIO) accounts have been advanced as an accounting framework to help measure environmental performance (Minx et al., 2009; Wiedmann, 2009). Using CBA to complement traditional assessments of environmental impacts opens a variety of new policy options for alleviating environmental pressures (Peters, 2010; Wiedmann and Barrett, 2013).

While there is a consensus on the basic approach that should be used to calculate CBA metrics – a Leontief demand-pull model (Leontief and Ford, 1970) (generally with use of monetary tables) – there has been less discussion about consensus on the actual values (Peters et al., 2012). Recent years have seen a proliferation of global MRIO tables that are used with standard Leontief models to calculate consumption-based footprints (Tukker and Dietzenbacher, 2013). While these accounts ostensibly seek to reach the same result – a global production and consumption database with explicit representation of trade – due to various implementation details there are nevertheless appreciable divergence between results as published by various research groups.

*Corresponding author. E-mail: daniel.moran@ntnu.no

With the limited success of Kyoto style accounting for controlling levels of greenhouse gas (GHG) emissions (Aichele and Felbermayr, 2012) policy-makers are beginning to turn to consumption-based approaches (Harris and Symons, 2013) to address issues related to carbon leakage (Peters, 2010). A strong concern of policy-makers is that the results that they are basing policy formation on are both consistent and robust (EU FP7, 2012). Hence, frameworks are required in order to compare CBA results across different models, and to provide a requisite understanding of variability between the results. With this in mind, we develop the use of uncertainty analysis within the goal of exploring model comparability and convergence.

In this paper we ask (a) How much do CBA results vary across current MRIO models? (b) Do CBA results of each model fall within the variance bounds defined by current estimates? (c) What are the contributions to variation and uncertainty of different parts of (generalised) MRIO systems, including the environmental satellite accounts, the description of global economic structure, and the description of demand?

1.1. Convergence between Models

Should we assume that there is a 'best' MRIO representation, and that results obtained from such an MRIO are near the average of, or at least bounded by, prior estimates? Conceptual differences to methods of analysis may occur, but at the most basic level the MRIO is still a collection of reconciled statistics (United Nations Department for Economic and Social Affairs Statistics Division, 1999).[1] Conceptual differences aside, it may be assumed that as the field advances, as data quality improves and as methods to reconcile data improve, that models will be attracted towards this correct statistical description of the world economy. However, each MRIO implementation suffers from some errors or differences in construction (see for an introduction, Wood et al., 2014). Different implementations are built with different target audiences and applications in mind. Builders must allocate scarce resources to the aspects of their model most salient for their intended purpose, though in doing so neglect other areas. If we can assume that besides conceptual differences, the impact of these choices relates to underlying data quality, and that the data quality is described stochastically (Lenzen et al., 2010), then we posit that across a set of MRIO implementations the variation due to actual stochastic errors should cancel. Hence as the sample size of MRIO models grows, after controlling for conceptual differences, we expect convergence of common results through increasing error cancellation. In a general sense one may hope, without ever being able to prove analytically, that continued improvements in modelling will increase convergence towards the underlying correct statistical account and that convergence of results is better than divergence. This convergence is not necessarily uniform; one implementation may be better in all ways than others. Measuring each observation's distance from an average value is only one indicator of how much confidence may be placed in that observation. Further, given a consistent set of modelling choices applied to the statistical account, it is then possible to analyse policy-relevant issues, such as GHG emissions embodied in final consumption that reflect the MRIO construction and not the application.

[1] At the statistical level MRIOs are represented as Supply and Use Tables, and MRIOs embody a modelling choice already, but we more loosely use the term MRIO to refer to collection of databases relevant for CBA.

Hence, we pose this as our first research question: for each country, how convergent are the results of CBA emission estimates based on different MRIOs?

1.2. Variance Between Models Due to Differences Between the Stressors

Each MRIO's environmental CBA result can be understood simplistically as a product of three variables: a flow matrix **Z** describing the economic structure, an environmental stressors matrix (or 'satellite account') **F** describing the per-sector direct environmental impacts of production, and a consumption bundle **Y** describing the composition of final consumption. The total CBA footprint C is a function of these three variables: $C = f(\mathbf{F}, \mathbf{Z}, \mathbf{Y})$. Of these, we assume that economic structure **Z** generally has a higher uncertainty than **Y**, and we observe that the environmental stressor **F** often has the greatest variance across models.

Even for a basic GHG emissions stressor, there is still substantial room for variability: what precisely should be included in the inventory, which data source(s) should be used to construct the inventory, and how the total impact should be allocated amongst particular sectors, since GHG emission inventories are rarely available itemised in a manner compatible with the MRIO's economic sector classification (Marland, 2008). For GHG emissions there are differences between the models in how many GHGs are included, which emission sources are included/excluded, how sectoral inventories are estimated if empirical data are not available, and, if including non-CO_2 GHGs, how the gasses are characterised in terms of their global warming potential. Industrial process emissions, solvent and other product use are generally included in the more recent MRIO models. Agricultural and waste emissions are sometimes included, and land-use change and forestry emissions are generally not included due to the difficulty in establishing cause and effect mechanisms in an MRIO framework. Whilst fuel combustion emissions are essentially the simplest form of emission, strongly linked to specific economic activities, here we are still faced with variability across the models in regard to how cross-border flows of fuels are accounted for. Some of these cross-border flows relate to the impact of purchasers by residents abroad (particularly regarding motor vehicle transit), whilst other flows relate to the extent that international transport activities are included especially regarding the bunkering of fuels. These differences become more acute for stressors that are more difficult to measure or to allocate to particular economic sectors, e.g. land area or biodiversity impact (Stadler et al., 2014). It would greatly help if energy accounts were consistent with the System of National Accounts (United Nations Statistics Division, 1993), but there is a lack of data in this convention, with most energy (and hence fuel combustion emissions) organised according to energy balances (International Energy Agency, 2012) where model builders have to use a variety of assumptions, trade statistics, and transport statistics to convert from the territorial to residence principle and to allocate 'activity' data to an industry (and household) classification.

Differences in how these details are managed cause substantial variability in the stressors used by each MRIO model at the statistical level. Some of these issues are ignored, some issues are treated with simplistic assumptions, and some are treated with detailed bottom-up models that do not always agree with top-down estimates. Whilst such estimates influence MRIO reliability, the issues are not unique to MRIO modelling, and are problematic across the statistical community and for current climate policy needs. We feel that it is important to separate these problems, which could be conceived as conceptual bias – different ideas of what we want to measure and how to do it – in the environmental satellite account from the more generic stochastic uncertainty of conceptually equivalent estimates. The issues are important for understanding of MRIO results, but the data quality here could (and should) be addressed outside of the MRIO models.

There are fewer sources of conceptual bias between MRIO models in how to construct the economic flows matrix $\mathbf{Z}$. All current MRIO models seek to allocate production-based emissions to final consumers using the monetary flows linking producers and consumers as a proxy for the flow of embodied emissions. Monetary input–output (IO) tables are a well-studied subject with, compared to environmental satellite extensions, more established and standardised accounting practices, and fewer sources of conceptual bias. This is not to say that there is perfect agreement on how to construct IO tables. The relevant accounting standards, backed by the UN System of National Accounts, are evolving.

Owen et al. (2014) apply structural decomposition analysis (SDA) to several global MRIOs in order to separate the effects of differences in $\mathbf{Z}$ and $\mathbf{Y}$ between models. In SDA, constituent variables are held constant while others are allowed to change, allowing one to determine how influential each constituent variable is in determining the final result. In this study, we follow a similar idea by exogenising the effect of environmental stressors, $\mathbf{F}$. This will allow us to see how much of the variation the CBA footprint result, C, is due to differences between how each MRIO model describes the global economic structure and final consumption versus differences in the environmental stressors used in each. We hypothesise that much of the difference between footprint results will be explained by the differences in the environmental stressors matrix used by each MRIO builder.

If our hypothesis is correct, namely that the biggest source of difference between CBA results comes from differences in $\mathbf{F}$, it would suggest that the MRIO community should turn more attention to harmonising the stressors between accounts to ensure the stressor matrices measure the same things, in the same way, with the same line-item distinctions and sectoral allocations. This would be a comparatively easy step that could eliminate much of the disagreement between CBA results.

1.3. Variance Within Each Model Due to Stochastic Error

Recalling the previous definition of the total CBA footprint, C, as a function, $f(\mathbf{F}, \mathbf{Z}, \mathbf{Y})$, by using the same stressor, $\mathbf{F}$, we can remove bias in the stressor, isolating how much of the difference in the total footprint, C, is due to differences in the flow matrix and final demand matrices.

One approach used to estimating the internal reliability of the model results when faced with stochastic error is to use Monte Carlo (MC) analysis (Bullard and Sebald, 1988; Lenzen et al., 2010; Nansai et al., 2012; Wilting, 2012). Quandt (1958) proposed that the values in an IO table are not absolute, but merely point estimates within some probability distribution. Quandt's original work and recent work by Wilting (2012) assumed the errors were normally distributed, though others (Lenzen et al., 2010) have also assumed log-normal distributions. In this study we rely on West's (1983; 1986) finding that results are relatively insensitive to the functional form chosen, and in the absence of empirical data indicating otherwise, a normal distribution was chosen for simplicity. Thus, the value of each element in $\mathbf{Z}$, $\mathbf{F}$, and $\mathbf{Y}$ can be understood as the mean value (μZ_{ij}) of a normal distribution with some standard deviation $\sigma_{Z_{ij}}$. In MC analysis the formula $\dot{C} = (\dot{\mathbf{F}}, \dot{\mathbf{Z}}, \dot{\mathbf{Y}})$ is repeatedly solved for perturbed variables $\dot{\mathbf{Z}}, \dot{\mathbf{F}}$, etc., where $\dot{\mathbf{Z}}$ is sampled from the normal distribution $N(\mu\mathbf{Z}, \sigma_Z)$. The standard deviation of the population $\dot{C}$ can be taken as an estimate of its variance. The repeated perturbations simulate the construction of many MRIOs each with some small errors.

One critique of this approach is that it implicitly assumes that every variable (transaction), or more specifically, the variance of every variable, is independent (Wilting, 2012). If errors are correlated and not independent, a more refined MC approach would be required. By perturbing flows rather than coefficients, we remove a dependency between the variables, but it can still be expected that large energy flows are correlated between the stressor matrix and the flow matrix.

2. METHODS

We perform an MC analysis for six different MRIOs under six scenarios with various permutations of exogenised **F** and **Y** matrices and regimes for estimating standard deviations. The four global MRIO models compared are: EXIOBASE (Tukker et al., 2013), both at original 129-sector-per-country resolution ('EXIOBASE') and aggregated to 15 sectors per country ('EXIOBASE15'), WIOD (Dietzenbacher, Los et al., 2013), the OpenEU MRIO (Weinzettel et al., 2011; Galli et al., 2012) which is based on the GTAP database (Global Trade Analysis Project, 2008; Andrew and Peters, 2013), and Eora (Lenzen et al., 2012; 2013; Moran, 2013), again both at original resolution ('Eora') and at an aggregated 26-sector-per-country resolution ('Eora26').[2] All the MRIOs were provided as industry-by-industry IO tables (IIOT), with the exception of Eora which is a heterogeneous MRIO but is implicitly converted to an IIOT MRIO during the Leontief inversion (Lenzen and Rueda Cantuche, 2012). The procedures for exogenising the **F** and **Y**matrices and the various regimes for estimating the relative standard are described below.

In each scenario, we run an MC analysis for the standard environmentally extended Leontief model $C = \mathbf{s}(\mathbf{I} - \mathbf{A})^{-1}\mathbf{Y}$, where $\mathbf{x}$ is a vector of sectorwise gross output $\mathbf{x} = \mathbf{Ze} + \mathbf{Ye}$ ($\mathbf{e}$ is a column vector of 1s). The vector of emissions intensities $\mathbf{s}$ is $\mathbf{s} = \mathbf{f}\hat{\mathbf{x}}^{-1}$, where $\mathbf{f}$ is a vector of per-sector total environmental impact (here $\mathbf{f}$ is Gg (kt) CO_2 emissions from fossil fuel burning per sector, so $\mathbf{s}$ is CO_2/\$ of production in each sector). The technical coefficients matrix $\mathbf{A}$ is derived from a flows matrix $\mathbf{Z}$, which describes the inter-industry flows in monetary terms, normalised by gross output, thus $\mathbf{A} = \mathbf{Z}\hat{\mathbf{x}}^{-1}$. The result C is the total environmental emissions associated with final demand $\mathbf{Y}$. For the MC analysis, the elements in $\mathbf{Z}$ and $\mathbf{Y}$ are offset with a matrix of perturbations $\mathbf{E}^Z$ where each element e^Z_{ij} is sampled from the normal distribution $N(\mu z_{ij}, \sigma_{Z_{ij}})$, and similarly for $\mathbf{E}^Y$. Over repeated samples, the mean value of e^Z_{ij} will be zero. The Leontief system $\dot{C} = \dot{\mathbf{s}}(\mathbf{I} - \dot{\mathbf{A}})^{-1}\dot{\mathbf{Y}}$ is repeatedly solved with resampled perturbations so $\dot{\mathbf{Z}} = (\mathbf{Z} + \mathbf{E}^Z)$, $\dot{\mathbf{Y}} = (\mathbf{Y} + \mathbf{E}^Y)$, $\dot{\mathbf{x}} = \dot{\mathbf{Z}}\mathbf{e} + \dot{\mathbf{Y}}\mathbf{e}$, $\dot{\mathbf{A}} = (\dot{\mathbf{Z}})\hat{\dot{\mathbf{x}}}^{-1}$, and $\dot{\mathbf{S}} = (\dot{\mathbf{F}})\hat{\dot{\mathbf{x}}}^{-1}$, in order to obtain a population of $\dot{C}$ results.

The scenarios considered for each MRIO are listed in Table 1.

The methods for harmonising **F** and the regimes for estimating the relative standard error (RSE) (σ) will now be explained.

[2] Other models, such as GRAM based on the OECD IO tables (Wiebe et al., 2012), were not included because they were not publicly available to the authors at time of publication.

TABLE 1. The scenarios considered.

Scenario #	Harmonised **F**	σ Regime
1	No	$\sigma_F = 0.1; \sigma_Z = 0.1; \sigma_Y = 0.1$
2	No	Logarithmic
3	Yes	Logarithmic
4	Yes	$\sigma_F = 0.1; \sigma_Z = 0.1; \sigma_Y = 0.3$
5	Yes	$\sigma_F = 0.2; \sigma_Z = 0.3; \sigma_Y = 0.1$

2.1. Harmonising the Satellite Accounts

To construct Scenario 3, the total value of the environmental stressor was exogenised by using the same total value for each country across the MRIOs. The stressor investigated was CO_2 emissions from fossil fuel burning (the line item detail of included emissions sources is noted in the Supplementary Information). The total territorial CO_2 emissions for each country were taken from Eora (the choice of the baseline value is irrelevant to the results, but Eora provides a superset of year and country coverage) and were allocated sectorwise using the allocation pattern each MRIO originally used for its CO_2 satellite account.

2.2. Setting a Common Year

The MRIOs investigated do not all cover the same time period. Eora covers 1990–2011, WIOD covers 1995–2009, EXIOBASE (first version) covers just 2000, and OpenEU covers just 2004. In order to allow direct comparison in a single year, the Eora CBA results were used as a proxy in order to re-scale the EXIOBASE territorial emissions forward from 2000 to 2002 and the OpenEU territorial emission results backwards from 2004 to 2002. Economic data are left unscaled to avoid balancing issues. This step was only needed for analysis requiring a common base year. The median change in production emissions was +4% between 2000 and 2002, and −8% between 2004 and 2002. While ideally adjusting the results to a common comparison year would not be necessary, we argue the scaling is a necessary convenience and is not so disruptive since it maintains the relative difference of the prior (or post) year for the common year, so that large divergence in results do not grow or diminish.

Each MRIO covers a slightly different set of countries so we only considered the subset of 41 countries covered by all the MRIOs.

2.3. Regimes for Estimating RSE (σ)

An MC analysis was run using two different regimes for estimating standard deviations. In the simple regime we assume the relative standard deviation of each element is 10% of its value, as per Wilting (2012). This provides a baseline result in Scenario 1. In all scenarios, following Rypdal and Winiwarter (2001) and Winiwarter and Rypdal (2001), we assume the relative standard deviation of each element of the environmental satellite account to be 10% of its value ($\sigma_F = 0.1$).

The second regime for estimating RSE assumes that standard deviations follow a power law distribution. Previous research has shown that variance is not homogeneous. Larger values tend to be more accurate, because stochastic errors cancel (Quandt, 1958; Lenzen et al., 2010). Our regime for estimating each element's standard deviation is based on Quandt's (1958) finding that errors follow a logarithmic distribution (standard deviations decrease exponentially as the value increases), and here we use the regression of UK IO table elements from (Lenzen et al., 2010) and say $\sigma_{Z_{ij}} = 0.393|z_{ij}|^{-0.302}$ for all elements in the MRIO table. Since the MRIO accounts are in units of million EUR, a €1 billion transaction ($\sigma_{Z_{ij}} = 1 \times 10^3$) will be estimated to have a relative standard deviation of $\sigma_{Z_{ij}} = 4.8\%$ and a €1 million transaction will have an estimated relative standard deviation of 39%. Clearly, the estimates of RSE from the UK will not be applicable globally. However, we feel the choice is justified because we are interested mostly in showing the impact of assuming every element has an RSE of 10% to assume that the RSEs follow a logarithmic distribution, and because the UK estimates were very conservative.

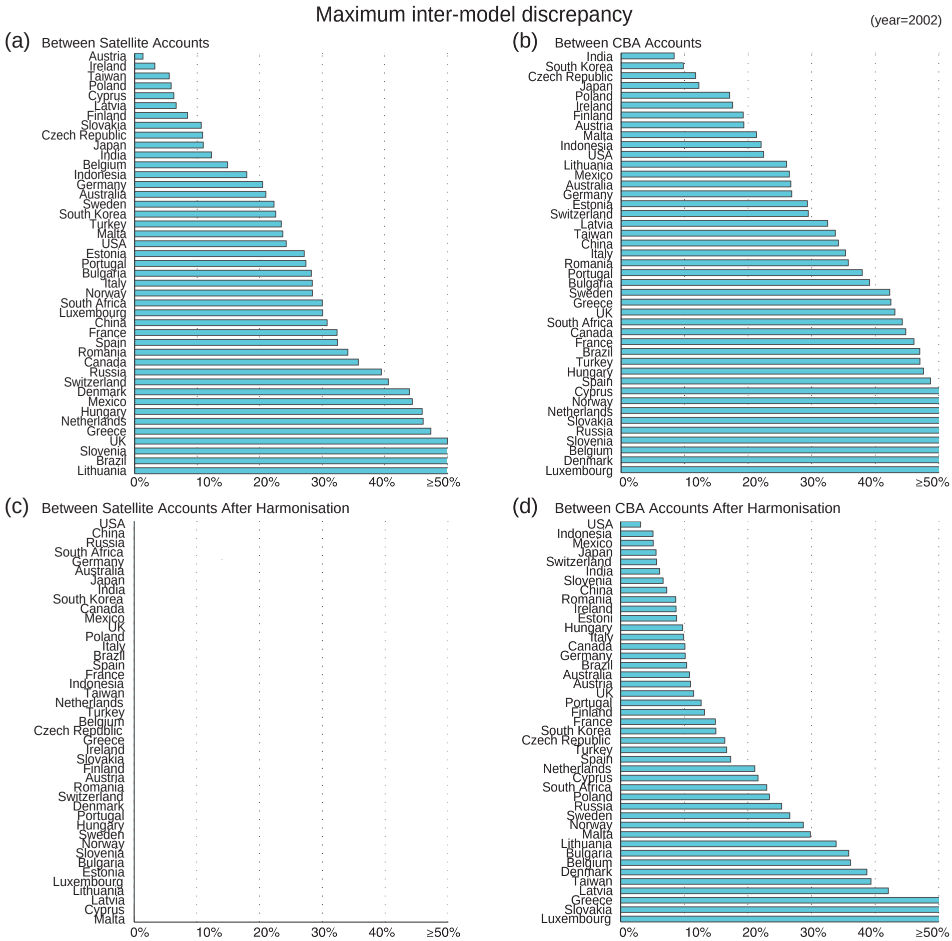
2.4. Handling Imbalances in the MC Perturbations

For an IO table to be valid it must be balanced. Each sector's gross output must equal the sum of its inputs. The perturbed MRIOs will not satisfy this fundamental balancing condition, but are compensating for the imbalance by calculating gross output on the perturbed sums of intermediate Z and final y demand (ignoring the difference in column sums). It is not computationally feasible to re-balance the MRIOs after each perturbation (this would require tens of thousands of hours of compute time). Furthermore, there are various different methods of balancing (Jackson and Murray, 2004) which do not give a common unique solution, and we are again confronted with the question of interdependence of matrix elements. Thus, we take the CBA results from the perturbed MRIOs as is.

3. RESULTS

A first indication of model behaviour may be obtained by plotting the maximum amount of disagreement between CBA results for each country relative to the smallest CBA model result. This gives a 'worse-case' assessment of potential disagreement across all MRIO models. In reality, this is often due to one model being the outlier across the set. We plot four figures, Figure 1(a) shows the disagreement in the production accounts from the MRIO databases as taken. This is the difference in the 'raw' environmental stressor. Generally, we see maximum difference of less than 20% in this account, in line with known differences in emission data sets (Rypdal and Winiwarter, 2001; Marland, 2008). Some examples of large differences are Luxembourg where WIOD is an outlier from all other models, possibly due to different methods to treat residential to territorial principles, as cross-border trade of citizens is common here. Denmark, in comparison, has weak consensus between the models (though we note this result could be due to inconsistent accounting of bunker fuels between models). In Figure 1(b), we see the impact of the application of the demand-based model, with maximum discrepancy shown for the consumption-based results. Here, the impact of differences in economic structure and final demand also play a role. As expected, maximum difference between the models increases when including the effects of structure and demand to the stressor. The magnitude is in the range of 5–10% increase for most countries. Figure 1(c) simply illustrates that in Scenario 3, the production account is harmonised across all models. Figure 1(d) then shows the maximum discrepancy of consumption-based results from a harmonised stressor total (the allocation can still differ

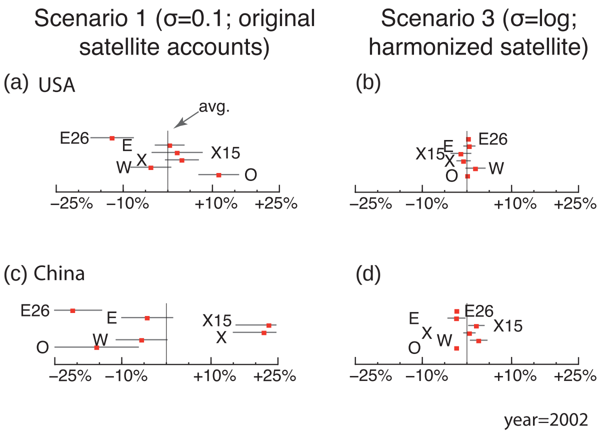
FIGURE 1. Largest disagreement over territorial emissions and CBA results between the MRIOs, before (Scenario 1) and after (Scenario 3) satellite account harmonisation.



amongst models). Most countries fall in a 20% or less discrepancy range, with about a third of the countries in 5% or less. For the USA, China, and India, and many other major economies no two models provide CBA results that differ by >10% in this last scenario. For Germany, Japan, and UK, the biggest disagreement is $\approx 8-12\%$. South Africa, Russia, and Sweden are three larger economies with larger disagreements between MRIOs, 16% and 24%, respectively.

It is perhaps more insightful to look beneath this maximum divergence at individual model results, and see if we get clustering around a model mean, and whether the results lie in the range of our predicted data uncertainty. Figure 2 plots the model clustering for the USA and China before (left-hand column) and after (right-hand column) harmonising the satellite account. In this figure, CBA results from each model are plotted as the relative distance from the multi-model mean (vertical centreline). The error bars denote the range of one standard deviation as determined from the MC analysis. Prior to the satellite harmonisation, the

FIGURE 2. CBA carbon footprint results from each of the MRIOs (year 2002, EXIOPOL and OpenEU interpolated) relative to the multi-model mean (vertical centreline).



Note: Left-hand panels (a) and (c) show results from Scenario 1 (where elements are assigned 10% RSE) and right-hand panels (b) and (d) show results from Scenario 3 (where satellite accounts have been harmonised and elements are assigned log-distributed RSEs). Results convergence improves dramatically after harmonising the satellite account.

variation between models is up to ±10% for the USA and nearly ±25% for China. After the harmonisation step, the relative disagreement between models is <5% for both. This implies that for these two large economies, despite being quite trade exposed, we see convergence within 2 if not 1 standard deviation across the models – clearly the models are agreeing on flow through of impacts through the economy, despite being of highly varying levels of detail (Eora uses the natural classification of some 400 sectors for the USA, whereas WIOD is at the other extreme with 35 sectors). Aggregation for these two countries is actually shown to have a limited impact within models, with aggregated versions of Eora and EXIOPOL producing results in agreement with their disaggregated versions.

Figure 3 presents the results from Scenario 3, with harmonised satellite accounts for all 43 countries covered in common by the MRIOs. As can be seen, for most major economies (Germany, Japan, France, and India) the relative distance from the mean is generally less than 10%. Russia provides an interesting case. As noted above, even after satellite harmonisation, in Scenario 3, the difference between the smallest and largest CBA result is greater than 30%. Yet no one model is >15% away from the multi-model mean. This underscores the fact that the choice of measure (exactly how one compares the various MRIO results against each other), and framing of results (for example, naming the chart 'model divergence' versus naming it 'model convergence') can have a big impact on the perceived reliability of MRIO results. Figure 3 (following Figure 2) also presents the summary results for the MC analysis, showing the error propagation through the model. From these results, it is clear that even with a harmonised stressor, our choice of estimate for the logarithmic regression equation was too optimistic given the variation in model results. Even for the UK, we obtain model results that do not agree within the first standard deviation. For larger countries, it appears that our error estimates were reasonable (we see model agreement within one standard deviation), whereas for many smaller countries there is clearly a much higher variation between models than what we anticipate from our error estimates. This raises the question

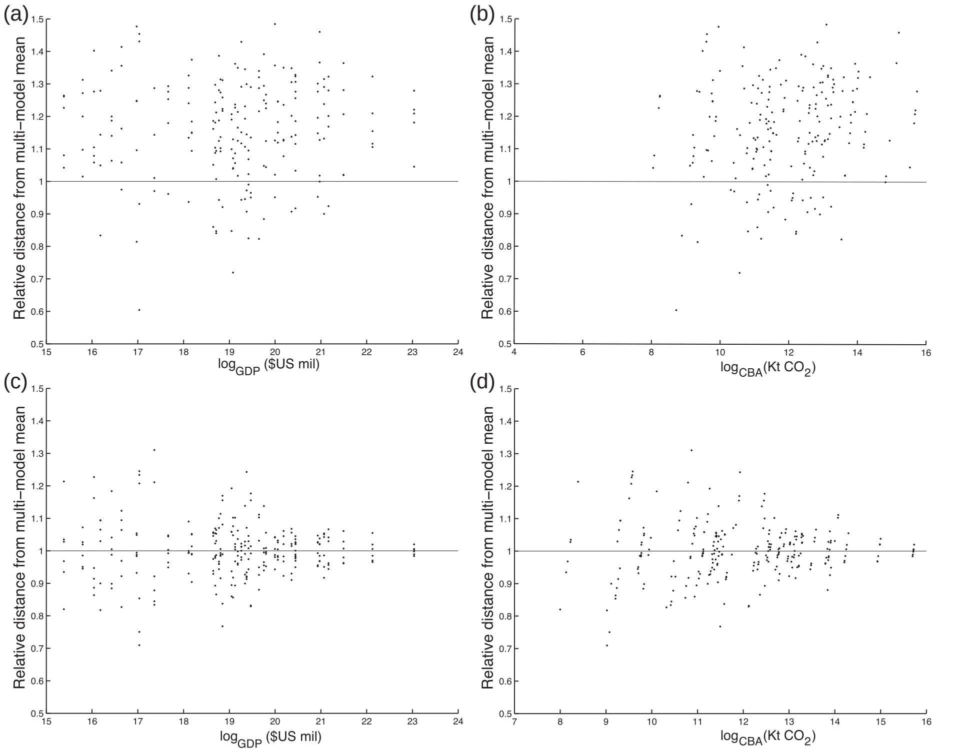
FIGURE 3. Countrywise CBA carbon footprint results from each of the four MRIOs relative to the multi-model mean (vertical centreline).

Note: Error bars indicate one standard deviation of the results, as determined from the MC analysis. Convergence for the USA and China is good, while variance for Sweden, Russia, and South Africa is lager. Results shown are from Scenario 3 (i.e. log-normal RSE; harmonised satellite accounts). Data are for year 2002.

of what ‘anticipated’ source data errors we would need in order for our models to agree. We return to this question later, before exploring the other interesting insight from these results: that there is a relationship between accuracy and country size.

We thus investigate whether model convergence improves for countries with larger GDPs or CBA carbon footprints. Figure 4 shows the relative distance between each country’s CBA result as determined by each MRIO and the multi-model mean, against GDP (left-hand panel) and absolute CBA value (right-hand panel). In Scenario 1 (top panels, un-harmonised stressor), the relative distance from the multi-model mean has only a weak decrease with larger values of GDP or emissions. On harmonising the stressor (Scenario 3, bottom panels, harmonised stressor), we see a much stronger correlation. This supports the theory that large economies are relatively well studied, at least in economic terms (intermediate and final

FIGURE 4. Model convergence as a function of GDP (panels a and c), and territorial emissions (panels b and d).

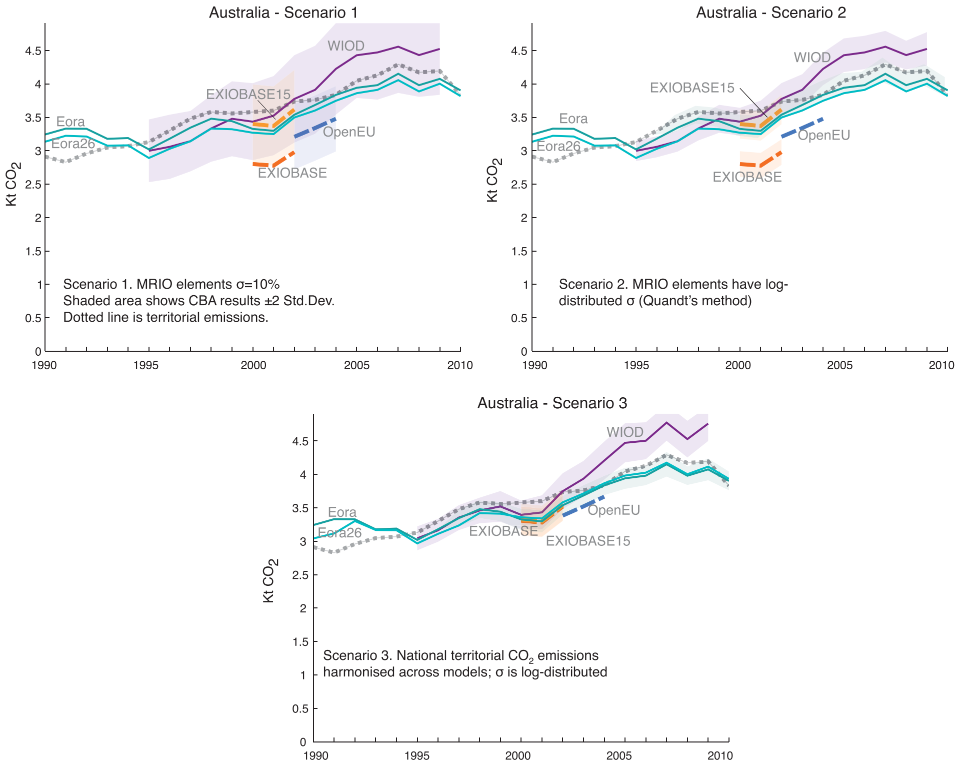


Note: Points indicate individual country CBA results as calculated by the various models. In Scenario 1 (top row) convergence does not improve for countries with larger GDP (left) or CBA carbon footprints (right). In Scenario 3 (bottom row), with the harmonised satellite accounts, convergence improves for the larger economies and emitters.

demand flows are well known). The difference between the top row (Scenario 1) and the bottom row (Scenario 3) implies that some work needs to be done to bring environmental accounting up to the robustness of economic accounting.

We can also look at results for individual countries. Figure 5 shows the time series of CBA results for Australia from each model under each scenario. We selected Australia as a representative country, covered by all four MRIOs. Full results for all countries are available in the online Supplementary Information and at http://worldmrio.com/comparison. In Scenario 1 the CBA results are found using the territorial CO_2 emissions account provided by each MRIO. The shaded two standard deviation area shows the range of perturbed CBA results from the MC analysis where all elements are assigned an RSE of 10%. The vertical bar shows one standard deviation of the results. Scenario 2 is the same except the standard error of elements in the MRIO was assigned using a power distribution relative to transaction size as described in Section 2.2. In this scenario, the results are still divergent but the reported confidence is much higher than with the larger standard error assumed in Scenario 1. It is interesting to see just how large the reduction in uncertainty becomes using the regressed RSE values. It is clear that model builders, whilst necessarily ensuring complete coverage

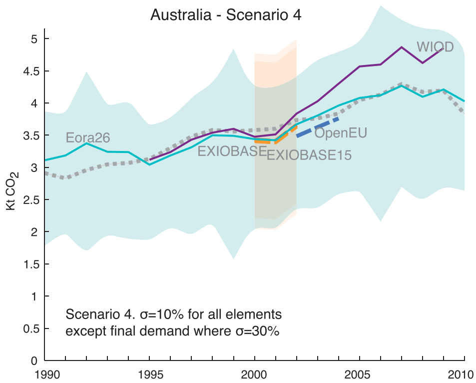
FIGURE 5. CBA carbon footprint time series for Australia under Scenarios 1–3.



Note: Data and charts for all countries available in the Supplementary Information and online at online at http://worldmrio.com/comparison.

of the economy, should focus their efforts on only the largest transactions in order to get convergent results, a conclusion argued previously by Jaynes (1957). For a system the size of EXIOBASE, it is less than 0.5% of transactions (typically) that are greater than 1 billion euro and hence have a significantly lower regressed RSE of source data (see Section 2.2). The level of (dis)aggregation can affect this outcome as well, particularly in the Eora MRIO. In the Eora MRIO countries use heterogeneous classifications and larger developed economies generally have more disaggregated IO table classifications or Supply-Use tables rather than simple IO tables. This means that the larger economies may tend, due to more disaggregation, to have smaller individual transaction values. In Scenario 3 (right-hand panel) the standard deviations are again assigned using a power distribution as in Scenario 2, but the total territorial emissions for all countries has been harmonised to the value reported by Eora so that the stressor is no longer a source of divergence in the results. Despite this harmonisation, we still observe divergence in the results. As the error bounds in both Scenarios 2 and 3 essentially capture the expected variability due to stochastic differences in the **A** and **Y** matrices of the models, and if we can control for the magnitude of differences in the stressor as per Scenario 3, then it follows that the observed differences in the model outcomes in Scenario 3 are then due to either differences in allocation of the stressors to individual sectors, or uncertainty in the values of the stressors themselves.

FIGURE 6. CBA carbon footprint time series for Australia under Scenarios 4 and 5.



Note: Data and charts for all countries available in the Supplementary Information and online at online at http://worldmrio.com/comparison.

In Figure 3, we reflected on the fact that Scenarios 1 and 3 estimates of source data uncertainty were too optimistic – models did not converge within one or even two standard deviations for all but the biggest economies. Now we explore Scenarios 4 and 5 (Figure 6) which explore how sensitive the CBA results are to larger perturbations of the technical coefficients and final demand values. In Scenario 4 all the intermediate demand values are assigned an RSE of 10% and the final demand is assigned an RSE of 30% ($\sigma_Z = 0.1$; $\sigma_Y = 0.3$). In Scenario 5, we have the inverse: intermediate flows are assigned an RSE of 30% and final demand elements an RSE of 10%. Both scenarios use the harmonised satellite accounts established for Scenario 3. As seen in Scenario 4 panel in Figure 6, the carbon footprint is highly sensitive to larger perturbations of the final demand. But as seen in the next panel, Scenario 5, the carbon footprint is not nearly so sensitive to even large perturbations of the intermediate flows matrix in the MRIOs. The variance in Scenario 5 results, with $\sigma_Z = 0.3$ being similar to the variance in Scenario 1 results with the much smaller $\sigma_Z = 0.1$. This reconfirms Jaynes' hypothesis that many of the values in the intermediate flows matrix of an IO table can be substantially perturbed and result in only relatively minor changes in the Leontief multipliers (Jaynes, 1957).

4. DISCUSSION

Our results show that there is still substantial quantitative variation between models even after harmonising **F** (Scenarios 1 and 3). Our initial hypothesis was that the biggest source of difference between CBA results comes from differences in **F**. We have not found enough evidence to support this hypothesis. Even after harmonising the total of **F** between models, significant difference in CBA results remain. Roughly, this is in the range of a maximum 5–30% discrepancy per country between all models. This difference between model results is in many cases larger than the one standard deviation results interval established by the stochastic MC analysis.

Even in Scenarios 3 and 4, where the stressor variable is harmonised, we observe that the difference between model results is in many cases larger than the year-on-year change.

Despite these facts, the results for the temporal change across models appear to agree; that is, the difference between the models appears to be reasonably constant and not fluctuating yearly. We cannot, in the present study, verify this statistically, but it does provide an insight into the robustness if not the precision we are able to obtain from MRIO models – we may end up with quantitatively different results, but in general, we have qualitatively similar outcomes (Schoer et al., 2013).

As all MRIO models are essentially attempting to achieve the same thing in terms of consumption-based allocation of GHGs (the allocation of production-based emissions to consumers), ideally the research community would be able to provide harmonised databases for this to occur. A recent article led by guest editors of this journal discussed the vision and difficulties of creating ever-more comprehensive and coordinated MRIO models (Dietzenbacher, Lenzen et al., 2013). In the interim, we have found it useful to have a range of models in which to cross-check the validity of our results, and we see many of the major global efforts converging on harmonised results for at least the most well-known countries. At the present different MRIO models, for different reasons, publish a range of figures for territorial CO_2 emissions in each country; providing options to harmonise this situation is relatively straightforward and will substantially improve model convergence. Following the stressor harmonisation, our results show the relatively higher sensitivity of results to final demand than intermediate transactions (Figure 5). As a first step, trade data, as a significant portion of final demand, could come under scrutiny. Outside of the MRIO realm, it would be beneficial if Comtrade (UN, 2009) and associated databases were consistent with aggregate country statistics and UN aggregates.

4.1. Policy Application

Consumption-based accounting is gaining relevance as a potential policy tool. MRIO-based CBA accounts need to be stable and replicable if CBA is to be used for policy. Our initial results show that in CBA accounts' results there are disagreements up to 10–20% for major economies. The degree of agreement/disagreement varies by country and by which models are compared. The level of model agreement required depends on the application: a 15% disagreement in CBA results may be acceptable when investigating some research topics but unacceptably large for answering other questions. Generally, national emission targets are being set against a baseline, looking at the relative change across time. For the time-series data we have, models are generally showing consistent trends such that consumption-based emissions can be assessed as a counter-part to domestic emission measures. In terms of reaching an absolute goal of consumption-based emission levels, we feel that for most countries more consistency is required across the model, at least for legal recourse. The Supplementary Information file provides full numerical results from this study so users can decide for themselves whether the MRIO results they are using are subject to acceptable or unacceptable disagreement between models.

To some extent these disagreements are due to different definitions of the environmental stressor used. This source of disagreement will be comparatively easy for MRIO builders to reconcile. But the remaining disagreement due to the different descriptions of the economic structure, differences in the value and composition of final demand and trade, and differences in the sectoral allocation of GHGs emitted from production. These differences are more difficult to reconcile between MRIO models. Policy-makers need trustworthy tools. Efforts to explain and then reconcile differences between MRIO implementations will help towards this goal. In the meantime, it is important that MRIO users understand how much confidence may be placed in the results of MRIO analysis. Comparing model convergence is one method for measuring this reliability.

The online Supplementary Information file and the graphs at http://worldmrio.com/comparison provide countrywise results from our MRIO comparison exercise. The trustworthiness of each MRIO's CBA results can be assessed both by how sensistive those results are to perturbances in the technical coefficients matrix, and by how far they lie from the mean value of all CBA results from other MRIOs. We invite MRIO users to make use of these results to communicate how stable are the currently avaialable global MRIO accounts.

4.2. Further Research

In this paper, we have tried to analyse the sources and magnitude of variability between MRIO models with the end goal of identifying how to improve convergence between model results. More work needs to be done on providing additional variance information on source data used in both this and earlier work. In particular, transactions in the trade blocks could be considerably less certain than values in the domestic IO table blocks. Furthermore, we exogenise the impact of the environmental stressors in this work. This can be refined by separating magnitude and structural (allocation) effects of the stressor; and by including information on stochastic uncertainty of the final environmental data set. Several next steps for further research into this question present themselves. One step is to investigate how sensitive are carbon footprints not just to the total level of territorial emissions but to the sectorwise composition of those emissions. Clearly, how bunker fuels and transport emissions are allocated is of importance here. Another remaining question is to check how much variability there is in the carbon footprint results at a product level. In this study we have looked at variability and convergence at the national level, but not at the product level.

FUNDING

This work was supported by the European Union CARBON CAP project [grant number 603386].

SUPPLEMENTAL DATA

Supplemental material for this article is available via the supplemental tab on the article's online page at http://dx.doi.org/10.1080/09535314.2014.935298

References

Aichele, R. and G. Felbermayr (2012) Kyoto and the Carbon Footprint of Nations. *Journal of Environmental Economics and Management*, 63, 336–354.

Andrew, R.M. and G.P. Peters (2013) A Multi-Region Intput Table Based on the Global Trade Analysis Project Database (GTAP-MRIO). *Economic Systems Research*, 25, 99–121.

Bullard, C.W. and A.V. Sebald (1988) Monte Carlo Sensitivity Analysis of Input-Output Models. *The Review of Economics and Statistics*, LXX, 708–712.

Dietzenbacher, E., M. Lenzen, B. Los, D. Guan, M.L. Lahr, F. Sancho, S. Suh and C. Yang (2013) Input-Output Analysis: The Next 25 Years. *Economic Systems Research*, 25, 369–389.

Dietzenbacher, E., B. Los, R. Stehrer, M. Timmer and G. de Vries (2013) The Construction of World Input-Output Tables in the WIOD Project. *Economic Systems Research*, 25, 71–98.
EU FP7 (2012) Quantification of Consumption-Based Emissions of Greenhouse Gases and Assessment of Policy Options. EU, FP7-ENV-2013-two-stage.
Galli, A., J. Weinzettel, G. Cranston and E. Ercin (2012) A Footprint Family Extended MRIO Model to Support Europe's Transition to a One Planet Economy. *Science of the Total Environment*, 461–462, 813–818. doi:10.1016/j.scitotenv.2012.11.071
Global Trade Analysis Project (2008) GTAP 7 Data Base. West Lafayette, IN, Department of Agricultural Economics, Purdue University.
Harris, P.G. and J. Symons (2013) Norm Conflict in Climate Governance: Greenhouse Gas Accounting and the Problem of Consumption. *Global Environmental Politics*, 13, 9–29.
International Energy Agency (2012) Energy Balances. Online database at http://www.iea.org/statistics/topics/energybalances/, Accessed January 2014.
Jackson, R.W. and A.T. Murray (2004) Alternative Input-Output Matrix Updating Formulations. *Economic Systems Research*, 16, 135–148.
Jaynes, E.T. (1957) Information Theory and Statistical Mechanics. *Physical Review,* 106, 620–630.
Lenzen, M., K. Kanemoto, D. Moran and A. Geschke (2012) Mapping the Structure of the World Economy. *Environmental Science & Technology*, 46, 8374–8381.
Lenzen, M., D.D. Moran, K. Kanemoto and A. Geschke (2013) Building Eora: A Global Multi-Region Input-Output Database at High Country and Sector Resolution. *Economic Systems Research*, 25, 20–49.
Lenzen, M. and J.M. Rueda Cantuche (2012) A Note on the Use of Supply-Use Tables in Impact Analyses. *Statistics and Operations Research Transactions*, 36, 139–152.
Lenzen, M., R. Wood and T. Wiedmann (2010) Uncertainty Analysis for Multi-Region Input-Output Models – A Case Study of the UK's Carbon Footprint. *Economic Systems Research*, 22, 43–63.
Leontief, W. and D. Ford (1970) Environmental Repercussions and the Economic Structure: An Input–Output Aproach. *Review of Economics and Statistics*, 52, 262–271.
Marland, G. (2008) Uncertainties in Accounting for CO_2 from Fossil Fuels. *Journal of Industrial Ecology*, 12, 136–139.
Minx, J.C., T. Wiedmann, R. Wood, G.P. Peters, M. Lenzen, A. Owen, K. Scott, J. Barrett, K. Hubacek, G. Baiocchi, A. Paul, E. Dawkins, J. Briggs, D. Guan, S. Suh and F. Ackerman (2009) Input-output Analysis and Carbon Footprinting: An Overview of Applications. *Economic Systems Research,* 21, 187–216.
Moran, D. (2013) The Eora MRIO. In: J. Murray and M. Lenzen (eds.) *The Sustainability Practitioner's Guide to Multi-Regional Input-Output Analysis*. Champaign, IL, Common Ground, 63–68.
Nansai, K., Y. Kondo, S. Kagawa, S. Suh, K. Nakajima, R. Inaba and S. Tohno (2012) Estimates of Embodied Global Energy and Air-Emission Intensities of Japanese Products for Building a Japanese Input–Output Life Cycle Assessment Database with a Global System Boundary. *Environmental Science & Technology*, 46, 9146–9154.
Owen, A., K. Steen-Olsen, J. Barrett, T. Wiedman and M. Lenzen (2014) A Structural Decomposition Approach to Comparing MRIO Databases. *Economic Systems Research*, 26, 257–278.
Peters, G.P. (2010) Policy Update: Managing Carbon Leakage. *Carbon Management*, 1, 35–37.
Peters, G.P., S.J. Davis and R. Andrew (2012) A Synthesis of Carbon in International Trade. *Biogeosciences*, 9, 3247–3276.
Quandt, R.E. (1958) Probabilistic Errors in the Leontief System. *Naval Research Logistics Quarterly*, 5, 155–170.
Rypdal, K. and W. Winiwarter (2001) Uncertainties in Greenhouse Gas Emission Inventories — Evaluation, Comparability and Implications. *Environmental Science & Policy*, 4, 107–116.
Schoer, K., R. Wood, I. Arto and J. Weinzettel (2013) Estimating Raw Material Equivalents on a Macro-Level: Comparison of Multi-Regional Input–Output Analysis and Hybrid LCI-IO. *Environmental Science and Technology*, 47, 14282–14289.
Stadler, K., R. Wood and K. Steen-Olsen (2014) The 'Rest of the World'– Estimating the Economic Structure of Missing Regions in Global MRIO Tables. *Economic Systems Research*. doi:10.1080/09535314.2014.936831
Tukker, A., A. de Koning, R. Wood, T. Hawkins, S. Lutter, J. Acosta, J.M. Rueda Cantuche, M. Bouwmeester, J. Oosterhaven, T. Drosdowski and J. Kuenen (2013) EXIOPOL – Development and Illustrative Analyses of a Detailed Global MR EE SUT/IOT. *Economic Systems Research*, 25, 50–70.
Tukker, A. and E. Dietzenbacher (2013) Global Multiregional Input–Output Frameworks: An Introduction and Outlook. *Economic Systems Research*, 25, 1–19.
UN (2009). UN comtrade – United Nations Commodity Trade Statistics Database. New York, United Nations Statistics Division, UNSD.
United Nations Department for Economic and Social Affairs Statistics Division (1999) *Handbook of Input-Output Table Compilation and Analysis*. New York, United Nations.
United Nations Statistics Division (1993) System of National Accounts 1993. New York, United Nations Statistics Division, UNSD.

Weinzettel, J., K. Steen-Olsen, A. Galli, G. Cranston, T.R. Hawkins, T. Wiedmann and E.G. Hertwich (2011) "Footprint Family Technical Report: Integration into MRIO Model." http://www.oneplaneteconomynetwork.org/resources/programme-documents/OPEN_EU_WP2_EC_Deliverable_Technical_Document.pdf.

West, G.R. (1983) Approximating the Moments and Distributions of Input-Output Multipliers (Eighth conference of the Australian and New Zealand section of the regional science association, Armidale, Australia).

West, G.R. (1986) A Stochastic Analysis of An Input-Output Model. *Econometrica*, 54, 363–374.

Wiebe, K.S., M. Bruckner, S. Giljum and C. Lutz (2012) Calculating Energy-Related CO_2 Emissions Embodied in International Trade Using a Global Input-Output Model. *Economic Systems Research*, 24, 113–139.

Wiedmann, T. (2009) A Review of Recent Multi-Region Input-Output Models Used for Consumption-Based Emission and Resource Accounting. *Ecological Economics*, 69, 211–222.

Wiedmann, T. and J. Barrett (2013) Policy-Relevant Applications of Environmentally Extended Mrio Databases – Experiences from the UK. *Economic Systems Research*, 25, 143–156.

Wilting, H.C. (2012) Sensitivity and Uncertainty Analysis in MRIO Modelling; Some Empirical Results with Regard to the Dutch Carbon Footprint. *Economic Systems Research*, 24, 141–171.

Winiwarter, W. and K. Rypdal (2001) Assessing the Uncertainty Associated with National Greenhouse Gas Emission Inventories: A Case Study for Austria. *Atmospheric Environment*, 35, 5425–5440.

Wood, R., T. Hawkins, E. Hertwich and A. Tukker (2014) Harmonizing National Input-Output Tables for Consumption Based Accounting – Experiences in Exiopol. *Economic Systems Research*, forthcoming.