UNIVERSITY *of York*

White Rose
university consortium
Universities of Leeds, Sheffield & York

eprints@whiterose.ac.uk
https://eprints.whiterose.ac.uk/

Accepted Manuscript

Capturing the heterogeneity of sub-national production in global trade flows

Simon A. Croft, Christopher D. West, Jonathan M.H. Green
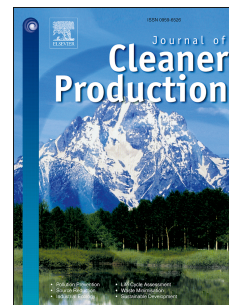
Please cite this article as: Croft SA, West CD, Green JMH, Capturing the heterogeneity of sub-national production in global trade flows, *Journal of Cleaner Production* (2018), doi: 10.1016/j.jclepro.2018.08.267.

Word count: 8000

Capturing the heterogeneity of sub-national production in global trade flows

Simon A. Croft [a*], *Christopher D. West*[a], *Jonathan M.H. Green*[a]

October 2017

*Corresponding author: simon.croft@york.ac.uk

[a] Stockholm Environment Institute
University of York
York
YO10 5DD
United Kingdom

Abstract

With increasingly complex and globalised supply chains, agricultural production and related impacts are often far removed from the point of final demand and difficult to trace. Accurately linking consumption to production is essential to understand drivers, key actors, and to facilitate actionable adaptation strategies to minimise negative impacts and guarantee food security. Here a hybridised multiregional input-output (MRIO) model, IOTA, is introduced. IOTA utilises sub-national and national level production, trade and environmental data, national scale commodity-use data, and a global economic MRIO, to link sub-national production and associated impacts to regional final consumption. In an example case-study, applying the model to Brazilian soy production and related land use for EU consumption, the relative levels of production in Brazilian states to meet EU demand differ from those of total production, and differ further still between the EU's constituent countries. Patterns can also vary considerably within a country's consumption profile depending on the sector of purchase. The linking of consumption to sub-national production and trade allows for more accurate and meaningful connections to be made between consumer behaviour and the associated impacts and risks. This enhanced understanding of consumption-driven impacts in turn informs, and allows for, more targeted and effective policy interventions to tackle the pressures and risks associated with agricultural commodity production for a global market.

Key words: Hybridised-MRIO; Supply chains; Production and consumption; Impacts; Food security; Teleconnection;

1 Introduction

Mapping and quantifying the production pressures and impacts driven by consumption activities allows understanding the demand drivers and actors involved in global supply chains; a vital precursor to the development of sustainable production-to-consumption systems (Godar et al., 2015). However, global supply chain networks are increasingly complex and the environmental impacts associated with commodity consumption often occur in distant countries (Kissinger and Rees, 2010; Lenzen et al., 2012a; Godar et al., 2015; Baldwin and Lopez-Gonazalez, 2015). Consequently, the linkages between consumption activities and the negative environmental impacts they drive become increasingly obscured (Lorek and Spangenberg, 2014; Spaargaren and Mol, 2008). The methods in this paper seek to overcome some of the biggest obstacles in connecting fine-scale production to final consumption activities and, in turn, allow for more accurate and useful linking of environmental impacts associated with production, via supply chain networks, to end consumers.

1.1 Material flow accounting

An intuitive approach to tracking supply chain paths is to use empirical data on production, trade and use; negating as much as possible the need for introducing modeling assumptions. However, for economic activities at national scales, such approaches - typified in Material Flow Accounting (Firscher-Kowalski et al., 2011) - are not usually sufficient for accurately making the connections between

2

source of production and destination of final consumption. Standard trade data typically provide a connection between origin of production and countries of first import, but fail to account for the remainder of the supply chain that links those imports through to final consumption. Whilst it is possible to incorporate data on some derived or secondary commodities to track subsequent stages of processed goods (Kastner et al., 2011), for many commodities this is a far from comprehensive approach to capturing all flows. For individual supply chains, process life cycle assessment approaches are typically used and can provide greater detail, but are time-consuming and the outcomes are specific and bespoke; limiting their ability to inform alternative supply chains or sourcing decisions (Bruckner et al., 2015).

1.2 Input-output models

Multiregional input-output (MRIO) models are an alternative method commonly used to overcome these limitations. MRIOs can represent the full supply chain from production to consumption (e.g. Wiedmann et al.,2015; Bruckner et al., 2015), employing monetary transaction data in the form of input-output (IO) tables (Miller and Blair, 2009). These IO tables detail monetary expenditure within an economy in matrix form, with rows detailing sector outputs as monetary value of sales, and columns showing the value of sector inputs as purchases. This includes intra- and inter-industry purchases, final demand (i.e. purchases from industrial sectors for final consumption), imports and exports, as well as value added (taxes, capital investment etc.).

An MRIO model is a combination of multiple national/regional-level IO tables, with the aggregated import and export data in the constituent IO tables extended to detail inter-industry and final demand purchases between different regions. MRIO tables capture the entire global economy (i.e. they account for all economic inputs and outputs); in capturing all monetary flows, the sales of goods and services implicitly capture data on all material flows in which a monetary transaction has occurred (Ibid). Standard IO methods allow for the expenditure data within an MRIO to be converted to determine input requirements per unit output for each sector (Leontief, 1936; Leontief, 1986). This means that for each unit of economic output a sector generates, the direct inputs from all other sectors can be determined. This information can in turn be used (via calculation of the Leontief inverse matrix) to determine indirect economy-wide dependencies which can be used to estimate embedded commodity flows (Ibid; see Section 2 for more details).

By "environmentally extending" sector-level production data to associated environmental impacts such as CO_2 emissions, land use, water use or biodiversity, production-driven impacts can be linked to consumption-driven demand via the economy-wide supply chains (e.g. Wiedmann, 2009; Lenzen et al., 2012b; Galli et al., 2012). Whilst a powerful tool, there exist restrictions which make it impossible within a standard MRIO to disentangle exactly what commodities are being produced, and precisely where, to satisfy regional demand.

1.2.1 Commodity resolution

A key limitation to the accuracy of commodity-level MRIO approaches to consumption based accounting is that the MRIO comprises sector-level data which

3

must be used as a proxy for individual commodity use within the economy. This means that two individual commodities which fall under the same sector classification within the MRIO (for example, soybean and oil palm might both be classified under an "oilseeds" sector) will be treated identically in terms of relative demand. If interest lies in a specific commodity from a consumption or production perspective, this presents an issue as different commodities within the same classification, such as soybean and oil palm, can have drastically varying uses and demand patterns, and an aggregation of the two returns a poor representation of either individually.

1.2.2 Spatial resolution

A major restriction on the potential utility of IO model outputs and associated impact indicators relates to the spatial resolution of the production which can be linked to consumption (Durán et al., in review). MRIOs typically operate at the country/regional level but national-level production data can present serious constraints when it comes to meaningful and appropriate pressure, risk or impact assessment. The extent to which, if at all, spatial resolution is problematic depends heavily on the commodity and producing country/region (i.e., the concentration of commodity production, and geographic extent of countries/regions, can vary significantly), and the specific impacts of interest. For example, where there is spatial heterogeneity in locations, quantities and/or resource intensities of production, the need for sufficient spatial resolution becomes important for assessing associated risks or impacts driven by supply chain activities. To illustrate, if water use is to be tied to water scarcity or the associated mitigation measures in place, or land use is to be linked to habitat loss and biodiversity impacts, it is important to make accurate connections between the consumer-system and the specific geographic context of production. Whilst national measures of some indicators can help set the political agenda, subnational information is often crucial for effective decision-making and targeted actions.

1.3 Overcoming limitations

By combining data in physical units of commodity production and supply with the monetary IO data (known as "hybridising"), MRIO models can be used to map the mass of individual commodities flowing through the different sectors and regions of the economy all the way from production to consumption (Giljum et al., 2008a; Giljum et al., 2008b; Ewing et al., 2012). The key goals of the methods described here are to retain individual commodity-level data as far as possible down respective supply chain paths, allocate these physical quantities of commodities as accurately as possible to respective regions and sectors, before utilising the sector-level expenditure data to complete the remaining paths at the consumption end of the supply chains. The "interface" between the physical and monetary datasets marks the point where individual commodity-level data joins sector-level data. As such, the further down the supply chain structure that individual commodity-level supply chain paths can be integrated, the more this aggregation effect can be alleviated.

Combining physical-monetary hybridisation with environmental extensions, models can be created which allow for commodity-specific production, and as-

sociated impacts, to be tracked. This combines the inherent benefit of IO approaches in capturing not only the direct requirements of consumers (i.e. purchases of the primary or directly derived commodities), but also indirect and embedded requirements (e.g. processed forms required by other industries to produce outputs, and inter-industry dependencies), with increased accuracy (Bruckner et al 2015). These methods provide a relatively simple, and broadly applicable, means to link consumption behaviour to (often) remote production and associated commodity-specific environmental impacts/risks.

1.4 IOTA model

This work presents a hybridised physical-monetary MRIO model, IOTA (Input-Output Trade Analysis). IOTA utilises commodity-level data in physical units and sectoral-level monetary expenditure data to map the entire producer to consumer supply chain for any given commodity for which data is available. National scale trade and commodity balance data are sourced from the Food and Agriculture Organization of the United Nations (FAO) which covers a number of agricultural commodities, whilst monetary MRIO data is sourced from the Global Trade Analysis Project (GTAP). Any MRIO could be used within the described framework, with GTAP chosen for its relatively disaggregated agricultural and regional coverage. Physical commodity-level data for production, trade, processing, waste and "use" are employed to more accurately capture and model the initial stages of the commodity-specific supply chains. This serves to extend the distribution of material flow into the monetary system and, thereby, more accurately capture some of the immediate processing steps of commodity-specific supply chains than aggregated IO tables would otherwise allow. This reduces the "aggregate sector" effect introduced by the IO tables, whilst retaining the considerable benefit of complete economic supply chain coverage they provide (Bruckner et al., 2015).

The key advancement is the utilisation of sub-national data at the production end of the supply chains, improving the spatial resolution of consumption-driven production mapping and, consequently, providing the opportunity for more accurate and meaningful impact assessment and evaluation. This information is fully integrated (via the retention of sub-national supply chain data, and not just an 'aggregation' of sub-national data). This is crucial for making meaningful links between consumption activities and on the ground impacts (Moran and Kanemoto, 2017; Wiedmann et al., 2011).

In Section 2 the IOTA model framework is outlined, with a specific focus on the novel incorporation of sub-national production and trade data and improvement of the physical-monetary data interface, which tackles two of the biggest limitations with conventional IO methods with an aim of improving resolution and accuracy. A comparison between the key component parts of the IOTA model and other notable supply chains models/techniques is provided in Figure 1.

In Sections 2 and 3 the model is then expanded and applied to a demonstrative case study of sub-national Brazilian soy production, and associated land use, embedded within European consumption (though the model can be applied to any commodity/producing region where data are available). Information provided by the model is explored further by focusing on soybean as a major source of animal feed; disaggregating sub-national consumption in Brazil
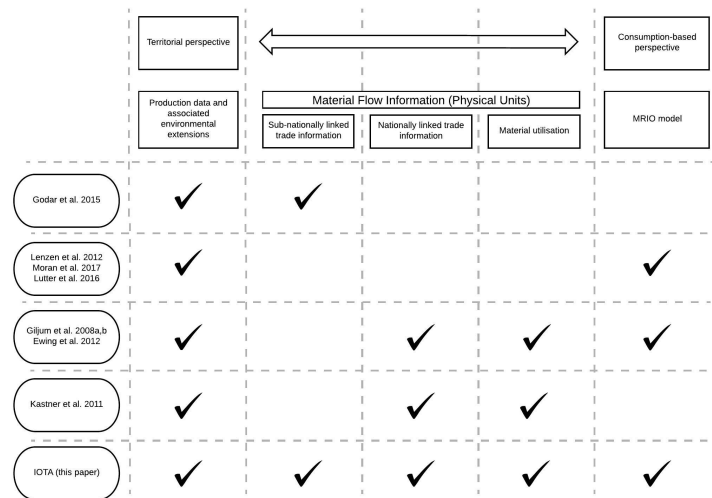
5

Figure 1: Visual representation of the spectrum of input data utilised within a variety of prominent supply chain models. Material flow information represents various information on commodity flows across supply chains, represented in physical units. IOTA seeks to enhance the linkages between production and consumption by incorporating information from multiple sources and stages of the supply chain paths.

that is embedded in Swedish final demand for meat and dairy products to illustrate the potential for heterogeneity in source region across key sectors of consumption.

This application demonstrates the utility of models that link sub-national production and trade to global consumption and, in turn, highlight the potential for such methods to provide policy makers and key supply chain actors with meaningful and actionable contextual information to tackle the drivers of production impacts such as land use change and natural habitat destruction.

2 Methods

2.1 Model overview

Within IOTA, different data sets are combined within a flexible framework in an attempt to provide an accurate mapping of producer to consumer supply chains whilst retaining broad scale (geographic and commodity) coverage. Where available, commodity-specific data is used to map the flow of a given commodity further through the system. When this data is not available (either for a given commodity, or for specific regions), a "next-best method" approach is taken. This ensures that broad scale coverage (spatial and commodity) remains, and allows for data sets providing partial coverage to still be utilised. In this paper the commodity choice of soy is used to demonstrate the model, but any commodity with available data could be implemented. Figure 2 provides a visualisation of the IOTA work-flow, and the following description gives a broad
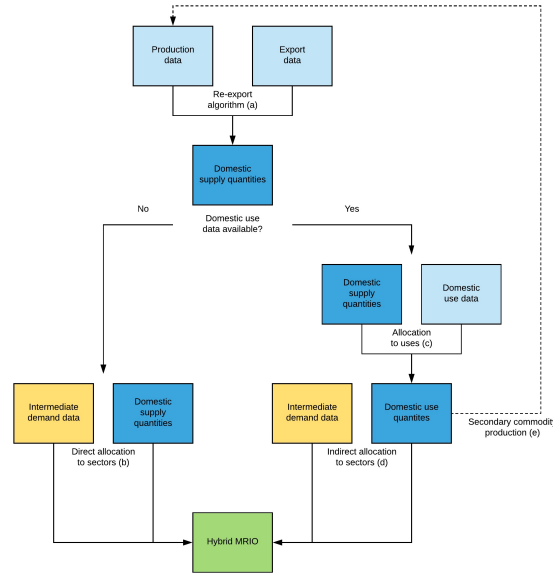
Figure 2: Figure 2: Simplified visualisation of model framework and data structure. Light blue denotes input data in physical units; dark blue calculated data in physical units; yellow input data in monetary units; green model data to apply. Production and trade data are used to obtain country-level domestic supply quantities via the re-export algorithm (a). Where use data are not available, relative expenditure within the MRIO (intermediate demand) is used to distribute domestic supply directly to economic sectors (b). Where use data are available, supply is distributed to use categories (c). Domestic use quantities are used in conjunction with the relative expenditure within the MRIO to more accurately distribute domestic supply to appropriate economic sectors (d). Where processing to secondary commodities does occur, this is fed back in as production of a "new" commodity at the start of the model framework (e). More detailed description of steps (a)-(e) provided in main text.

overview of how the different data sets and model components interact to form the model as a whole.

(a) Commodity-specific production and trade data is are run through a re-export algorithm (Appendix A) to link source of production to destination of import, removing re-exports. This provides each country with a "domestic supply" of a given commodity which retains an explicit link back to the point of production origin(s), as well as detailing where a country's exports end up. Note: the following description refers to country of origin for the sake of simplicity, but this can just as well be a sub-national region of production.

For example, this will provide the quantity of commodity supply available to Country A, how much of this supply originated in Country B, and how much (if any) of Country A's production went to Country C.

(b) If commodity-specific use-data *is not* available (either for the commodity in general, or for a specific country), the country-level domestic supplies are concorded (i.e. mapped and aggregated as appropriate) from FAO countries

to GTAP regions. This supply is then distributed amongst a region's economic sectors weighted according to relative use for different purposes and intermediate demand for the sector associated with the commodity's production (see below). From this point onwards, the monetary MRIO data is used to complete the supply chain.

For example, if Country A imports I tonnes of a commodity from Country B, all sectors within Country A have combined expenditure of J dollars on the global sector associated with the production of the commodity, and a given sector within Country A has expenditure K dollars on this global sector, then that sector within Country A will be allocated IK/J tonnes of Country A's supply imported from Country B.

(c) If commodity-specific use-data *is* available (for a given commodity and country), the country-level domestic supply is split proportionally between those uses in accordance with the relative use-data.

For example, if Country A imports I tonnes of a commodity from Country B, all Country A reported uses of this domestic supply have combined mass of J tonnes, and a given use is reported as utilising K tonnes, then that use will be allocated IK/J tonnes of the Country A commodity supply imported from Country B.

(d) The designated use quantities from (c) (excluding processing; see (e)) are concorded from FAO countries to GTAP regions. The quantities allocated to a specific use are then distributed amongst the associated GTAP sectors according to relative expenditure by these sectors on the appropriate supply sector. At this point the monetary MRIO data is used to complete the supply chain for all use other than "processing" (see (e)).

For example, if a given domestic use of a commodity in Country A is allocated I tonnes of Country B produced commodity from its domestic supply, all Country A sectors associated with food production have combined expenditure of J dollars on the global sector associated with the commodity production, and a given sector associated with the specific use has expenditure K dollars on this global sector, then that sector will be allocated IK/J tonnes of the Country A commodity supply imported from Country B.

(e) Domestic supply allocated to "processing" in stage (c) is converted to production of derived commodities. This is done proportionally according to the relative quantities of derived commodities reported as being produced (which are reported in raw material equivalence). These derived commodities are then run back through the model as separate commodities (see (a)) whilst retaining all information about the origin of the raw commodity.

For example, if Country A processes I tonnes of commodity sourced from Country B, and is reported to produce J and K tonnes of two directly derived commodities, the quantities of those commodities produced from the commodity sourced from Country B are estimated as $IJ/(J+K)$ and $IK/(J+K)$, respectively.

2.2 Monetary MRIO data

The monetary MRIO tables are built from the GTAP model following the methods described by Peters et al. (2011). Specifically, the GTAP9 database is used for reference year 2011 (the most recent reference year; previous reference years could be used), comprising 140 regions and 57 industrial sectors, and so 7980

$(140 * 57)$ individual economic entities. It should be noted that the general methods described here will be compatible with any other MRIO models also, subject to appropriate concordances. Following construction of the tables, the key components of interest are the "intermediate demand" matrix, Z, the "final demand" matrix, Y, and the derived "total output" vector, X.

Following standard IO methods (Leontief, 1936), the "technical coefficients" matrix, A, can be calculated as

$$A = Z\hat{X}^{-1}, \tag{1}$$

where $\hat{X}$ is a 7980x7980 diagonal matrix, with diagonal values $\hat{x}_{ii} = x_i$. The Leontief inverse, L, also known as the "total requirements" matrix, can then be calculated as

$$L = (I - A)^{-1}, \tag{2}$$

where I is the identity matrix. The total requirements matrix accounts for activities required across all sectors within the economy. Each element l_{ij} details the total outputs of sector i required per unit out by sector j.

While the technical coefficients matrix, A, allows for the calculation of direct output requirements needed to meet demands from certain sectors, the total requirements matrix, L, allows for the calculation of all outputs required across the whole economy to meet these same demands. This allows for consumption to be linked to production activities across the entire economy (sectoral and regional), and consequently to capture the embedded and indirect requirements common in today's complex, multi-actor supply chains.

2.3 National-level physical production, trade and use data

Commodity Balances data from FAO provide national-scale physical (typically mass) data for agricultural commodities detailing Domestic Supply (comprising information on production, imports, exports and stock variation), as well as uses of this supply for processing (production of secondary commodities), food (human consumption), feed (animal consumption), other uses, and waste. Combined with Detailed Trade Matrix data (annual trade totals, specifying trade partners; FAO), these data provide a national-level picture for 236 countries comprising their domestic supply, where this supply has come from, what this supply is used for, and where exported production has gone.

The Detailed Trade Matrix data need to be treated with care to avoid misinterpretation of the information they contain. Specifically, "re-exports" (the reported export of a previously imported commodity) need to be addressed to avoid incorrect allocation of commodity origin to an intermediate trade partner. The Netherlands, for example, is a major trade partner for much of Europe, and in 2011 are recorded as exporting nearly 1.5 million tonnes of soy (FAO). The Netherlands does not, however, domestically produce any soy; all of these "exports" are in fact re-exports of prior imports originating elsewhere. With 236 countries, it is not feasible to manually sort and adjust these data for all commodities to account for re-exports. Whilst re-exports can be dealt with easily via algebraic methods subject to assumptions about proportional distribution (e.g. Kastner et al., 2011), such an approach relies on balanced data (i.e. countries don't export more than they produce and import), which can require

adjusting production levels to suit. The FAO trade data do not balance in such a way, and require a method which not only reallocates traded commodities to their true origin, but also adjusts values to make the data balance.

An algorithm was designed to address data balancing and re-exports. The algorithm constrains total trade to total production, while country level exports are constrained by domestic production and imports. The result is a balanced set of production and trade data (i.e. total supply = total production) with trade allocated from country of origin (i.e. original production) to country of final import (the final destination of the raw good before processing/use). See Appendix A for more details.

2.4   Hybridising physical and monetary datasets

The merging of the physical and monetary datasets uses the calculated re-exports data to assign quantities of the physical commodity to the respective MRIO region of import. A simple concordance matrix is used to convert the 236x236 re-export matrix into a 236x140 format, detailing trade from the 236 FAO countries to the 140 GTAP regions. Distribution of the commodity within the importing region's economy (i.e. across the 57 GTAP sectors within each country) is performed according to the concorded FAO physical use data, and relative expenditure levels within the GTAP data (specifically the intermediate demand matrix, Z). For each of the 236 FAO countries where use data is available, these data is are used to calculate a weighted distribution of domestic supply which is utilised for food, feed, seed, processing, other uses, or is lost to waste. Waste is distributed back across the other uses proportional to relative weightings; in the absence of more detailed waste information, this method assumes a homogeneous level of waste as opposed to more complicated estimates, but importantly does ensure material equivalence is preserved and waste is not "lost" from the system.

All of the remaining uses are assigned to the most relevant sectors within the GTAP data base (with any unassigned sectors being allocated to "other uses"), with the exception of processing which is handled differently (see Section 2.5). For the rest of these uses, the relative expenditure of their associated sectors in the producing sector of the region of import is used to split the physical quantity of the commodity amongst them. For FAO countries where the re-exports data show domestic supply of the raw commodity but no use data is present, the relative expenditure of sectors (in the associated GTAP region(s)) on the producing sector is used to distribute the physical commodity across the region's economy, i.e. it is assumed that within a region, sector use is relative to sector expenditure. For the primary commodity, and each of the derived commodities, a 236x7980 physical-to-monetary matrix, $\tilde{Z}$, is produced, where $\tilde{z}_{ij}$ is the physical quantity of commodity originating from country i which is allocated to sector j (where "sector" refers to one of the 57 industrial sectors within one of the 140 regions present within the GTAP database).

2.5   Processing use data and derived commodities

Processing is typically associated with the production of secondary (or derived) commodities. In the case of soybean, this is the processing of soybean to produce soybean oil and soybean cake. As well as production, trade and use data

for the primary soybean commodity, FAO also provide these data for a number of secondary commodities. This means that the quantity of a country's domestic supply utilised for processing can be converted into production of these secondary commodities, and the first stage of the model run again to reallocate this production via commodity specific trade data rather than the sector-level monetary data within the MRIO. This is achieved by taking a country's allocation of domestic supply for processing, splitting it proportionally according to the relative levels of secondary products recorded as being produced, and then treating this production of secondary commodities in the same way as the initial production of the raw primary commodity.

It is worth stressing that the origin of the primary product is retained all the way through this process, meaning that the production of secondary commodities in a given country can still be linked to the original source of the primary commodity. The use data for these secondary commodities is then utilised as described in Section 2.4 for the primary product, with the exception that secondary products utilised for processing are now distributed across the appropriate MRIO sectors rather than being treated separately. If similar data for tertiary, or further order, commodities are available, this process could be extended as appropriate.

2.6 Environmentally extending the model

The above methods describe the means of establishing a link between the physical and monetary data to connect final demand to physical production quantities. To understand impacts related with consumption driven production, the model is "extended" via environmental coefficients. These take the simple form of a 236 element vector, $\vec{E}$, which provide a country-level per-unit-production value for a given commodity, for example area of land used per unit mass of soybean produced. Multiplying the physical production quantities required to meet demand by such an environmental coefficients vector converts the unit mass of production required into the associated input/impact.

2.7 Sub-national production and trade data

Sub-national level production, trade and land use data from the Transparency for Sustainable Economies platform (Trase, 2017; Godar et al., 2015) and Brazilian Institute of Geography and Statistics (IBGE, 2017) are employed to improve spatial resolution in the focal country of Brazil. In the results presented here, this is done at the state level, breaking down the national scale picture into 27 constituent jurisdictions. The data sets utilised describe production quantities of the primary commodity by state, along with import and export reports for primary and secondary commodities, linking each state to the country of export/import. Due to the use of different sources, available sub-national physical data is not fully equivalent to the national-scale data:

- Production data is available only for the primary commodity (soybean) and not the secondary commodities (oil and cake).

- Trade data covers state-level imports and exports from Brazil, but does not cover inter-state trade.

- Use data is not available at the state level.

Where data is consistent in form with that at the national scale, methods follow closely with those presented above for the national-level data. Where they differ, national-scale data is used where possible, assuming national-scale homogeneity. Given a more intimate knowledge and understanding of these sub-national data than the national scale data provided by FAO (which comprise of multiple sources), the trade data for sub-national Brazil is considered more reliable than the production estimates. As such, rather than let the re-export algorithm restrict trade to match reported production levels, instead Brazilian sub-national production is scaled, where necessary, to meet reported exports before running the re-export algorithm (note: this scaling is only required for one state). As there is no inter-state trade data available, this means that imports and exports from a given state are limited purely to the "domestic" production of that state, and imported goods previously imported to that state (as opposed to production from and/or imports to other states). In the case of Brazilian soybean, this is not likely to have a notable effect on results given the relatively low levels of imports (and thus low potential for re-exports). The individual states then behave within the context of the re-export algorithm just as any other country.

A lack of state-level data for use and the production of secondary commodities is tackled by utilising national-level data. For use classification and quantification, the national-level proportions of domestic supply reportedly used for each use are applied at the state-level to determine an estimate of the quantities allocated to processing (i.e. the production of secondary commodities) and the other uses. As with primary commodity production, processing data is then scaled to meet export requirements if necessary. As secondary commodities are often produced via the same processing activities (e.g. soybean oil and cake are both byproducts of the processing of the same soybean), where appropriate their production is linked according to national scale ratios of co-production. As the MRIO only contains national-level data, other uses are aggregated and homogeneously allocated across the national-scale economy as with national scale data. The estimated state-level production of secondary commodities is then used in conjunction with the state- and national-level data within the re-export algorithm to redistribute this production throughout the economy. Again, sub-national supply of secondary commodities is aggregated to the national scale and merged with the monetary data as previously described.

The final outputs for the model with sub-national data incorporated retain similar form, but with the producing country of Brazil replaced by its 27 constituent states, with full state-level linkages maintained. It is worth noting that that MRIO data remains at the original regional scale, i.e. Brazil remains at the national scale within the monetary data, and in the outputs demand for sectors within the Brazilian economy, and demand by Brazil itself, are still resolved at this scale. Figure 5 (Appendix B) provides a visualisation of the full flow chart for IOTA as implemented for Brazilian soy here, illustrating all modelling steps and the interactions of different input and intermediate data sets.
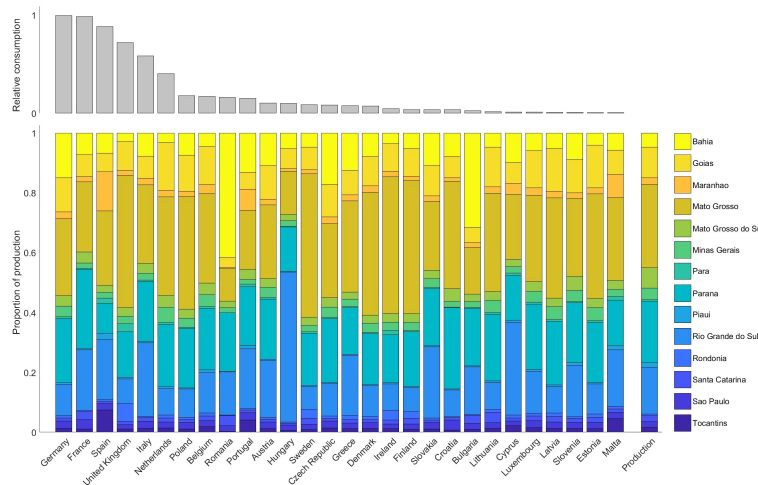
Figure 3: Brazilian state-level soybean production (for significant producing states) for EU demand, by country. Each bar details proportion of country's total Brazilian soy consumption fulfilled by individual states. The proportional contribution of each state to total Brazilian production is also represented in the right-hand bar. Relative total consumption of Brazilian soy per EU country is represented by the grey bars at the top of the figure.

3 Results

3.1 National scale

In 2011, global production of soy - used as an input to the IOTA model - totalled 262 million tonnes. Brazil was the largest soy producer, with production totalling nearly 75 million tonnes; over 28% of global production. It was also the biggest exporter of soy, with over 49 million tonnes of raw soy, and directly derived oil and cake, recorded as being exported. IOTA results show that the EU's total consumption of soy (that is consumption of raw, processed and embedded soy) in 2011 exceeded 39 million tonnes, approximately 15% of global production. Of this, 16 million tonnes (41%) originated in Brazil. The top three consumers of Brazilian soy within Europe (Germany, France and Spain, in descending order) accounted for nearly half (47%) of the EU's Brazilian soy consumption, whilst the ten lowest consuming countries accounted for just 3% collectively. Country-level demand for soy will vary depending on a number of factors, such as population size, wealth and diet.

3.2 Sub-national scale

Not only does total demand for soy vary significantly between different countries, but so too do their sourcing profiles and, consequently, the origins of production. Figure 3 visualises total demand by each EU country for Brazilian produced soy (upper bars), with relative supply from each Brazilian state shown (lower bars). Whilst the same key producing states typically dominate supply, the variability in relative demand from the 27 different states of origin is clear.

13

State	% Bra Prod	% Bra-EU supply	Min. relative demand	Max. relative demand
Mato Grosso	28%	29%	Romania 11%	Sweden 28%
Parana	21%	19%	Spain 10%	Croatia 27%
Rio Grande Do Sul	16%	16%	Sweden 8%	Hungary 50%
Goias	10%	9%	Romania 3%	Netherlands 16%
Other states	26%	27%	Hungary 14%	Romania 51%

Table 1: Proportion of Brazilian produced soy for total EU consumption provided by key producing states, and the maximum and minimum proportions these contribute to individual EU countries' Brazilian sourced supply.

The top four producing states in Brazil account for three quarters of the country's total production, and are responsible for a similar proportion of the EU's Brazilian sourced supply. Table 1 details the proportion of the EU's Brazilian supply that these states account for and, for each state, the specific EU countries for which it makes up the smallest and the largest proportion of that country's total Brazilian soy supply.

The relative levels of Brazilian sub-national production largely correlate to the bulk of Brazilian sourced soy for EU supply (columns two and three, respectively, Table 1). However, as also demonstrated within Figure 3, the variation from the EU average, and Brazilian production profiles, among constituent countries is notable.

For example, Mato Grosso, the biggest soy producing state in Brazil, provided 30% of the EU's Brazilian soy, yet accounted for nearly half (48%) of Sweden's Brazilian supply. It follows that if relative demand is higher from some countries then it must also be lower from others and, likewise, if a country has relatively high demand for production in one state, it must have correspondingly low demand for another. This can be seen for Rio Grande Do Sul, the third biggest producer, which accounted for 16% of total EU supply, yet only 8% of Sweden's Brazilian supply. Conversely, Hungary received 50% of its Brazilian soy from this state; over three times the EU-wide average.

Total 2011 global production of soy for Swedish consumption is estimated at 525 thousand tonnes, or 0.20% of global soy production. Over 236 thousand tonnes of Swedish supply originated from Brazil, ranking it 13th of the EU28 countries in terms of Brazilian-sourced soy.

3.3 Sector level

Variety in origin of Brazilian sourced soy does not just differ between consuming countries, but also within countries by sectors of final consumption. One of the main drivers of soy production is through the use of soy as feed for cattle, pigs and poultry. As well as the (generally domestic) land requirements for grazing and shelter, the requirements for feed production (which can often be remote) need to be considered in land footprints associated with meat and dairy production. IOTA allows for purchases from the different GTAP sectors, and

from different regions, to be linked back to the origin of embedded soy, and associated impacts, at the subnational scale. Figure 4 shows a visualisation of the relative land use requirements, by state, for Brazilian sourced soy due to Swedish consumption of feed-related sectors, specifically: "Cattle Meat" (*Fresh or chilled meat and edible offal of cattle, sheep, goats, horses, asses, mules, and hinnies. Raw fats or grease from any animal or bird*), "Other Meat" (*Pig meat and offal. Preserves and preparations of meat, meat offal or blood, flours, meals and pellets of meat or inedible meat offal; greaves*) and "Milk" (*Dairy products*).

Mato Grosso is the biggest producing state of soy for total Swedish consumption, and also for soy consumption embedded within the three key feed-related sectors. Figure 4 shows that roughly the same proportion of land use associated with production for feed-related sectors (38%) is embedded within purchases from the Cattle Meat and the Milk (dairy) sectors, with purchases from the Other Meat sector accounting for 24%. In contrast, soy from Bahia embedded in feed-related sectors for Swedish consumption is dominated by Other Meat, which accounts for approximately two thirds of this supply. Cattle Meat and Milk (dairy) are approximately equally responsible for the remaining use.

4 Discussion

A key novelty of this paper is the linkage of sub-nationally derived trade flows to a MRIO framework. The methods and results presented here represent an initial use of sub-national production and trade data to link national-level consumption activities to sub-national production and land use. These sub-national results are not achieved by simply post-downscaling national-level production and trade data, but rather are calculated by fully integrating sub-national information into the model and calculations. Given available data, production could equally be linked to more specific indicators of environmental impact, such as land use change, deforestation, and biodiversity risk.

Moran and Kanemoto (2017) state that "Previous work has linked consumption and supply chains to biodiversity impacts but only at the country level. Biodiversity threats are often highly localized. Knowing that a given consumption demand drives biodiversity threat somewhere within a country is not enough information to act." Whilst - in their study - they look at biodiversity risks at the sub-national scale, it is done so within a context of national-level production and trade (i.e. no explicit link it made between where production is taking place sub-nationally, and the localised threats discussed). The methods in this paper provide a means to link spatially explicit production data to localised impacts and risks of this kind.

The results emphasise the key role that feed-related imports from major trade partners play in embedded soy consumption, and more importantly how the impacts of this consumption, in terms of land use for soy production, can be highly heterogeneous within producing countries (Figure 4). Not only do the results show that production, and associated land use, varies considerably at the sub-national scale within Brazil, but production is concentrated in the highly biodiverse and highly threatened Cerrado biome (Strassburg et al., 2017), highlighting the potential impacts associated with production (Figures 3 and 4). Whilst consumption-mapping activities such as this are not necessary to know that soy production in Brazil is concentrated in such key regions, it is
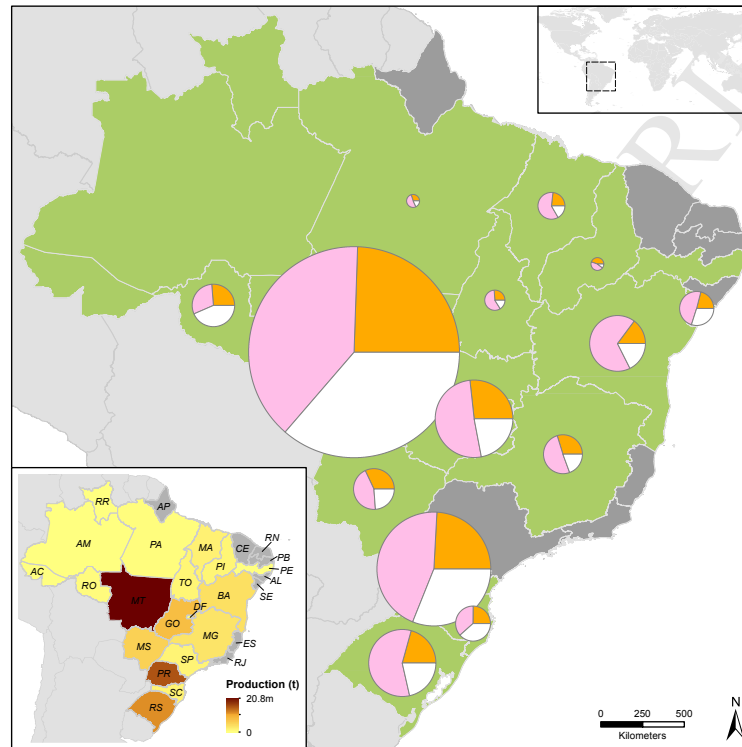
15

Figure 4: State-level land use for soybean production associated with Swedish consumption activity across three livestock-linked industrial sectors: cattle meat (orange); other meat (pink); and dairy (white). Radius of pie-charts indicates relative quantity of state land use (where it exceeds 100ha) for soy production embedded in Swedish final demand (range: 130ha in PI to 17,653ha in MT). Green/grey states denote production/no production takes place, respectively; inset shows total production intensity by state (AC: Acre; AL: Alagoas; AM: Amazonas; AP: Amapá; BA: Bahia; CE: Ceará; DF: Distrito Federal; ES: Esp¨ªrito Santo; GO: Goiás; MA: Maranhão; MG: Minas Gerais; MS: Mato Grosso do Sul; MT: Mato Grosso; PA: Pará; PB: Paraíba; PE: Pernambuco; PI: Piauí; PR: Paraná; RJ: Rio de Janeiro; RN: Rio Grande do Norte; RO: Rondônia; RR: Roraima; RS: Rio Grande do Sul; SC: Santa Catarina; SG: Sergipe; SP: S?o Paulo; TO: Tocantins).

16

important for explicitly linking direct and embedded consumption to these areas and providing information to help inform procurement decisions with an aim to minimising negative impacts abroad.

In contrast, traditional, national scale, impact assessments paint a homogeneous picture of production and impact patterns. For example, whilst European countries would provide different levels of demand and have different consumption profiles, the relative estimated production and impacts within a country such as Brazil would be the same for all consuming countries, simply scaled according to total consumption. Although post-downscaling of results would provide an approximation of sub-national distribution of production and impacts, it would still retain an identical relative sub-national impact profile for all consuming countries. The full integration of sub-national data into the production and trade mapping demonstrated in this work suggest this is far from the case, with Figure 3 highlighting the variation between different countries' consumption profiles, and how these differ from the Brazilian-wide production landscape. Whilst the example of land use is applied here as an environmental-extension, extensions could readily be applied for other impacts, and would be especially appropriate for other highly localised impacts (such as deforestation, habitat loss etc.).

Numerous countries and industry actors are signing up to agreements to avoid deforestation in supply chains (for example the New York Declaration on Forests; UN, 2014). Consumption-based approaches such as MRIO-based tools have the advantage of capturing both direct and indirect connections to risk "hotspots", but only with sub-national resolution of the kind presented here would individual nations or regions be able to delineate how their consumption profiles and associated supply chains contrast to other consumption profiles in term of linkages to areas which are highly heterogeneous in terms of their risk profiles. The results presented here illustrate how broadly adopted strategies and policies, for example EU-wide commitments to deforestation, require an understanding of how different consuming countries (and associated producing regions) contribute towards the problem. This isn't limited to just understanding the production profile of commodities like soy in Brazil and absolute consumption figures for consuming countries, but it is also about understanding the difference in sourcing and supply chain patterns.

This understanding is necessary for assessing the problem in the first instance, but also for acting effectively to adapt and mitigate against further environmental (and other) impacts. The variety in sourcing patterns, as well as consumption totals, suggests that individual countries would ideally adopt bespoke strategies to minimise their risks and impacts. Whilst broader scale policies might be useful, or indeed necessary, to enforce behaviour change, it is at a finer scale that interventions and appropriate strategies need to be enacted to most successfully and efficiently mitigate against, for example, increased destruction of natural habitat.

Within individual countries, it is demonstrated that sourcing patterns for purchases from different sectors can also be markedly different. This information allows for more precisely targeted adaptations in sourcing behaviour to minimise impacts and risk, with information now levelled at sector level purchases and sourcing rather than just at the national scale. The methods presented here also allow for exploration of purchasing patterns from different sectors from different countries. Differences between national- and sector-level sourcing patterns have

17

the potential to be enlightening from a perspective of domestics activities versus reliance on overseas industry, and again provide another piece of contextual information for targeted and effective adaptation activities.

Sweden is chosen as an exemplar due to its national commitment to its "Generational Goal" which explicitly recognises the potential for overseas impacts associated with consumption activities. The Generational Goal aims to "provide guidance on values that are to be protected and the changes in society that are needed if the desired quality of the environment is to be achieved" (Swedish EPA, 2017).

These methods and analysis are pertinent to this goal; both providing sub-national locational data on production quantities and land use (from which value-judgements can be made about areas that are linked to Swedish consumption which are worthy of protection), and providing details of the relative role of products of consumption (societal changes; i.e. cattle vs pork/poultry vs dairy consumption) that are linked to the production of soy in these regions.

In national implementation, policies and interventions are likely to vary depending on the sectors of consumption; for example, local dairy vs EU-wide meat. The former may necessitate changing sector-level policies for example, whilst the latter would require countries to hold bilateral discussions with other countries to reduce impacts. The data outputs from IOTA, such as those demonstrated here, have the potential to rapidly assess production/impact "hotspots", and how these might vary across supply and consumption contexts.

4.1 Assumptions and Limitations

Where it is necessary to make assumptions within the modelling framework when integrating sub-national scale and national scale data, homogeneity at the national level is implied to avoid introducing more complicated assumptions. Whilst there is clear divergence within the results, were more sub-national level data available (specifically information on the use of domestic supply and sub-national interactions at the inter- and intra-country level), the results would likely demonstrate even higher levels of heterogeneity than evidenced here.

Similarly, whilst the data used, and results shown, are at the state-level, this analysis could equally be run for higher resolution data; for example, municipality-level data is increasingly available and accessible (Godar et al., 2015). Just as increasing the resolution from national to state level results reveals a notable amount of sub-national variation, it is to be expected that moving from state level to municipality level analysis (a move from 27 Brazilian administrative units to over 5500) would in turn reveal sub-state variety, and with it yet greater sub-national heterogeneity.

The work here focuses on soy production in Brazil, but as more data becomes available this can just as well be applied to other commodities and producing regions. In addition, more of the detail stored in the datasets underpinning the sub-national trade data, such as importer and exporter information in bills of lading and customs declaration records (Trase, 2017), could be tied into these results to provide more contextual information to aid effective response and targeting of key supply chain actors. As producers, intermediaries and consumers push for increased transparency, this data is only likely to increase in availability and accessibility.

In the case of Brazilian soy as presented here, currently available data on sub-national trade is limited to commodity flows into and out of Brazil, and does not detail the movement of goods between regions within the country. Combined with a lack of sub-national data on the use of commodities, this means that flows of embedded goods (i.e. beyond the flow of raw soy and oil/cake derivatives) are treated homogeneously across Brazil. Consequently, beyond the tracked flows of the primary and secondary commodities, subsequent use and further processing is only modelled at the national scale. Again, this will have the effect of damping the heterogeneity in the system and, accordingly, the results.

The use of monetary MRIO data allows for completion of the supply chain, and extending the available data from point of import through to final consumption. However, this still operates at an aggregated sector level, and so is still susceptible to potential inaccuracies of aggregation artifacts. This is a trade off between absolute accuracy, and completion of supply chain mapping.

As additional data becomes available, the monetary data can be relied on less to complete the supply chain (i.e. empirical data can be employed to map the commodity flows yet further along the supply chains). Further, when monetary data is needed, they can be employed more accurately by increasing the accuracy of the sector allocations and merging of the two data types. For example, employing information on trading partners could allow for more accurate linking if imports to respective sectors, rather than assigning based on national level use data as currently occurs. Likewise, caloric information that details nutritional requirements of different livestock could be utilised to allocate soy used for feed across appropriate sectors within national economies (Kastner, 2011). Not only could such information improve the accuracy of the linking mechanics, but additionally it would increase the utility of final outputs for enacting appropriate policy interventions and strategies.

4.2 Conclusions

The methods and results presented offer great potential to link sub-national environmental and social data to dramatically improve the credibility of driver-to-effect linkages. Moran and Kanemoto (2017) stated that: "Improved spatial data for the trade model are especially important for spatially extensive countries such as the United States, China, Russia and India, where one industry may have different impacts across its domain". The results shown here highlight the variability that occur within production, impacts and supply patterns. Sub-national heterogeneity in consumption-driven production evident in the model outputs are even more crucial to understand when looking beyond, for example, land use, and applying this to contextual information such as deforestation or biodiversity impacts.

The framework is flexible enough to be applied to any commodities or regions where data are available, and provides scope for improvements in accuracy and resolution as the availability of sub-national data continues to improve (Pettorelli et al., 2014). With expansion of the production-linked indicators, and increased supply chain detail, a more holistic approach to impact and risk assessment can be undertaken (Green et al., in prep.). This offers the opportunity to overcome some of the biggest issues with footprint indicators to date; moving beyond broad-scale, generic risks which ultimately fail to provide much information at scales and levels of detail where effective interventions might take

place.

Acknowledgements

References

[1] Baldwin, R., Lopez-Gonzalez, J., 2015. Supply-Chain Trade: A Portrait of Global Patterns and Several Testable Hypotheses. The World Economy 38(11), 1682-1721.

[2] Bruckner, M., Fischer, G., Tramberend, S., Giljum, S., 2015. Measuring Telecouplings in the Global Land System: A Review and Comparative Evaluation of Land Footprint Accounting Methods. Ecological Economics 114, 11-21.

[3] Durán, A.P., Green, M.H., West, C.D., Visconti, P., Burgess, N.D., Virah-Sawmy, M., Balmford, A., in review. Putting species back on the map: devising a robust method for quantifying the biodiversity impacts of commodity-driven land conversion. Global Change Biology.

[4] Ewing, B.R., Hawkins, T.R., Wiedmann, T.O., Galli, A., Ercin, A.E., Weinzettel, J., Steen-Olsen, K., 2012. Integrating ecological and water footprint accounting in a multi-regional input–output framework. Ecological Indicators 23, 1-8.

[5] FAO. Food and Agriculture Organization of the United Nations. http://www.fao.org/faostat/en/#data

[6] Fischer-Kowalski, M., Krausmann, F., Giljum, S., Lutter, S., Mayer, A., Bringezu, S., Moriguchi, Y., Schütz, H., Schandl, H., Weisz, H., 2011. Methodology and Indicators of Economy-wide Material Flow Accounting. Journal of Industrial Ecology 15(6), 855-876.

[7] Galli, A., Wiedmann, T., Ercin, E., Knoblauch, D., Ewing, B., Giljum, S., 2012. Integrating Ecological, Carbon and Water Footprint into a "Footprint Family? of indicators: Definition and role in tracking human pressure on the planet. Ecological Indicators 16, 100-112.

[8] Giljum, S., Lutz, C., Jungnitz, A., Bruckner, M., Hinterberger, F., 2008. Global dimensions of European natural resource use: First results from the Global Resource Accounting Model (GRAM). Vienna, Sustainable Europe Research Institute.

[9] Giljum, S., Lutz, C., Jungnitz, A., 2008. The Global Resource Accounting Model (GRAM): a methodological concept paper. Vienna, Sustainable Europe Research Institute.

[10] Godar, J., Persson, U.M., Tizado, E.J., Meyfroidt, P., 2015. Towards more accurate and policy relevant footprint analyses: Tracing fine-scale socio-environmental impacts of production to consumption. Ecological Economics 112, 25-35.

[11] Green, J.M.H., Croft, S.A., Durán, A.P., Balmford, A.P., Burgess, N.D., Fick, S., Gardner, T.A., Godar, J., Suavet, C., Virah-Sawmy, M., Young, L.E., West, C.D. (in prep.) Linking global drivers of agricultural trade to on-the-ground impacts on biodiversity.

[12] https://www.gtap.agecon.purdue.edu/

[13] Brazilian Institute of Geography and Statistics (Instituto Brasileiro de Geografia e Estatística). https://www.ibge.gov.br/

[14] Kastner, K., Kastner, M., Nonhebel, S., 2011. Tracing distant environmental impacts of agricultural products from a consumer perspective. Ecological Economics 70, 1032-1040.

[15] Kissinger, M., Rees, W.E., 2010. An Interregional Ecological Approach for Modelling Sustainability in a Globalizing World—Reviewing Existing Approaches and Emerging Directions. Ecological Modelling 221(21), 2615-2623.

[16] Lenzen, M., Kanemoto, K., Moran, D., Geschke, A., 2012a. Mapping the Structure of the World Economy. Environmental Science & Technology. 46(15), 8374-8381.

[17] Lenzen, M., Moran, D., Kanemoto, K., Foran, B., Lobefaro, L., Geschke, A., 2012b. International Trade Drives Biodiversity Threats in Developing Nations. Nature 486, 109-112.

[18] Leontief, W.W., 1936. Quantitative input-output relations in the economic system. Rev. Econ. Stat. 18, 105-125.

[19] Leontief, W.W., 1986. Input-Output Economics (2nd Ed). Oxford University Press.

[20] Lorek, S., Spangenberg, J.H., 2014. Sustainable consumption within a sustainable economy - beyond green growth and green economies. Journal of Cleaner Production 63, 33-44.

[21] Lutter, S., Pfister, S., Giljum., 2016. Spatially explicit assessment of water embodied in European trade: A product-level multi-regional input-output analysis. Global Environmental Change 38, 171-182.

[22] Miller, R.E., Blair, P.D., 2009. Input-Output Analysis: Foundations And Extensions. Cambridge University Press

[23] Moran, D., Kanemoto, K., 2017. Identifying species threat hotspots from global supply chains biodiversity. Nature Ecology & Evolution 1.

[24] Peters, G.P., Andrew, R., Lennox, J., 2011. Constructing an environmentally extended multiregional input-output table using the GTAP database. Economic Systems Research 23, 131-152.

[25] Pettorelli, N., Laurance, W.F., O'Brien, T.G., Wegmann, M., Nagendra, H., Turner, W., 2014. Satellite remote sensing for applied ecologists: opportunities and challenges. Journal of Applied Ecology 51(4), 839-848.

[26] Spaargen, G., Mol, A.P.J., 2008. Greening global consumption: Redefining politics and authority. Global Environmental Change 18(3), 350-359.

[27] Strassburg, B.B.N., Brooks, T., Feltran-Barbieri, R., Iribarrem, A., Crouzeilles, R., Loyola, R., Latawiec, A.E., Oliveira Filho, F.J.B., Scaramuzza, C.A. de M., Scarano, F.R., Soares-Filho, B., Balmford, A., 2017. Moment of truth for the Cerrado hotspot. Nature Ecology & Evolution 1.

[28] Swedish EPA, 2017. http://www.swedishepa.se/Environmental-objectives-and-cooperation/Swedens-environmental-objectives/The-generational-goal/

[29] Trase, 2017. www.trase.earth

[30] UN, 2014. Forests: Action Statements and Action Plans, Climate Summit 2014. UN Headquarters, New York. (Accessed at: https://www.un.org/climatechange/summit/wp-content/uploads/sites/2/2014/07/New-York-Declaration-on-Forest-%E2%80%93-Action-Statement-and-Action-Plan.pdf)

[31] 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.

[32] Wiedmann, T., Wilting, H.C., Lenzen, M., Lutter, S., Palm, V., 2011. Quo Vadis MRIO? Methodological, data and institutional requirements for multi-region input-output analysis. Ecological Economics 70(11), 1937-1945.

[33] Wiedmann, T.O., Schandl, H., Lenzen, M., Moran, D., Suh, S., West, J., Kanemoto, K., 2015. The Material Footprint of Nations. Proceedings of the National Academy of Sciences 112(2), 6271-6276.

A   Re-export algorithm

For a given commodity, country level production can be considered as a vector, P, listing production quantity (e.g. by mass) for n countries, such that

$$P = \begin{bmatrix} p_1 \\ p_2 \\ \vdots \\ p_n \end{bmatrix}. \tag{3}$$

Similarly, exports can be considered as a matrix, E, where e_{ij} denotes the export from country i to country j, such that

$$E = \begin{bmatrix} 0 & e_{12} & \dots & e_{1n} \\ e_{21} & 0 & \dots & e_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ e_{n1} & e_{n2} & \dots & 0 \end{bmatrix}. \tag{4}$$

The diagonal entries are all zero valued since a country does not "export" to itself; the quantity of domestically produced goods that remain within the producing country will instead be derived from production and export totals.

Combining production and trade data allows for domestic supply, D, to be calculated with information of origin. Domestic supply is simply defined as production plus imports minus exports. Here d_{ij} is the quantity of country j's domestic supply which originates from country i.

$$D = \begin{bmatrix} d_{11} & d_{12} & \dots & d_{1n} \\ d_{21} & d_{22} & \dots & d_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ d_{n1} & d_{n2} & \dots & d_{nn} \end{bmatrix}, \tag{5}$$

where d_{ii} is domestic supply of domestically produced goods, that is

$$d_{ii} = p_i - \sum_{j \neq i} d_{ij}. \tag{6}$$

If D is constructed directly from available production and trade data, the issue of re-exports manifests itself in the form of erroneous allocation of traded commodities to the wrong country of origin. Whilst in some instances such allocations can be effectively invisible (i.e. there is no obvious indication that an incorrect allocation of production has occurred), in others it creates a clear issue to overcome as the calculated domestic supply of domestically produced goods, d_{ii}, can be negative.

Using reported production, P, and trade, E, data, the algorithm calculates domestic supply, D. The algorithm works by repeatedly performing a two-step operation for N iterations (here $N = 10,000$). Initially D is a zero matrix, and the two-step operation is as follows:

Step 1. $1/N$ of each country's annual domestic production is added to their domestic supply (i.e. $d_{ii} = d_{ii} + p(i)/N$).

Step 2. $1/N$ of annual exports are moved from each country's domestic supply to the appropriate recipients. These exports are proportionally comprised of domestically produced and imported goods according to current domestic supply, and are capped by available domestic supply.

The first time Step 2 is performed, domestic supply is purely domestically produced as no trade has yet been modeled. But the domestic supply soon becomes comprised of both domestically produced and imported goods, and subsequent calculations of trade account for exports and re-exports, whilst never allowing a country to trade more than it possesses. The outcomes of this are that:

- Total global domestic supply is constrained and equal to total global production (i.e. total quantity of goods is conserved).

- No country exports more than it produces, or re-exports more than it imports (i.e. no negative domestic supply values).

- Commodity origin is linked back to the source of production, not intermediate traders (i.e. trade links true origin to final destination).

As such, rows in D will be zero-valued for non-producing countries, and the domestic supply for each country (columns) will only contain non-zero entries for countries where production has occurred and with who trade (direct and/or indirect) has taken place. A copy of the core code for the re-export algorithm can be found in Listing 1.

Listing 1: MATLAB code for core re-export algorithm

```
1  \label{lst:re_exports}
2  % number of iterations
3  N=10000;
4
5  % load Production vector
6  load('P.mat');
7  % load Export matrix
8  load('E.mat');
9
10  % number of countries
11  CtryNo=numel(P);
12  % pre-calculate diagonal Production matrix
13  Pd=diag(P);
14  % pre-allocate Domestic Supply matrix
15  D=zeros(CtryNo);
16
17  for n=1:N;
18      % STEP 1: allocate production
19      % allocate production to domestic supply
20      D=D+Pd/N;
21
22      % STEP 2: perform trade
23      % calculate proportions of domestic supply required
   for each component of export iteration
24      temp1=E/N./(repmat(sum(D),CtryNo,1))';
25      % sum to check if greater than 1 (if domestic supply
   is less than desired export total)
26      temp2=repmat(sum(temp1, 2),1,CtryNo);
27      % constrain export greater than domestic supply to be
    equal to domestic supply
28      temp1(repmat(sum(temp1,2)>1,1,CtryNo))=temp1(repmat(
   sum(temp1,2)>1,1,CtryNo))./temp2(repmat(sum(temp1,2)
   >1,1,CtryNo));
29
30      % proportional change in domestic supply
```

24

```
31      e_n=ones(CtryNo,1)−sum(temp1,2);
32      % apply to domestic supply of domestic production (
        non−traded component)
33      e_n=diag(e_n)+temp1;
34      % take care of 0/0 cases
35      e_n(isnan(e_n))=0;
36      % take care of x/0 cases
37      e_n(isinf(e_n))=0;
38
39      % rescale domestic supply to redistribute according
        to trade
40      D=D∗e_n;
41 end
```
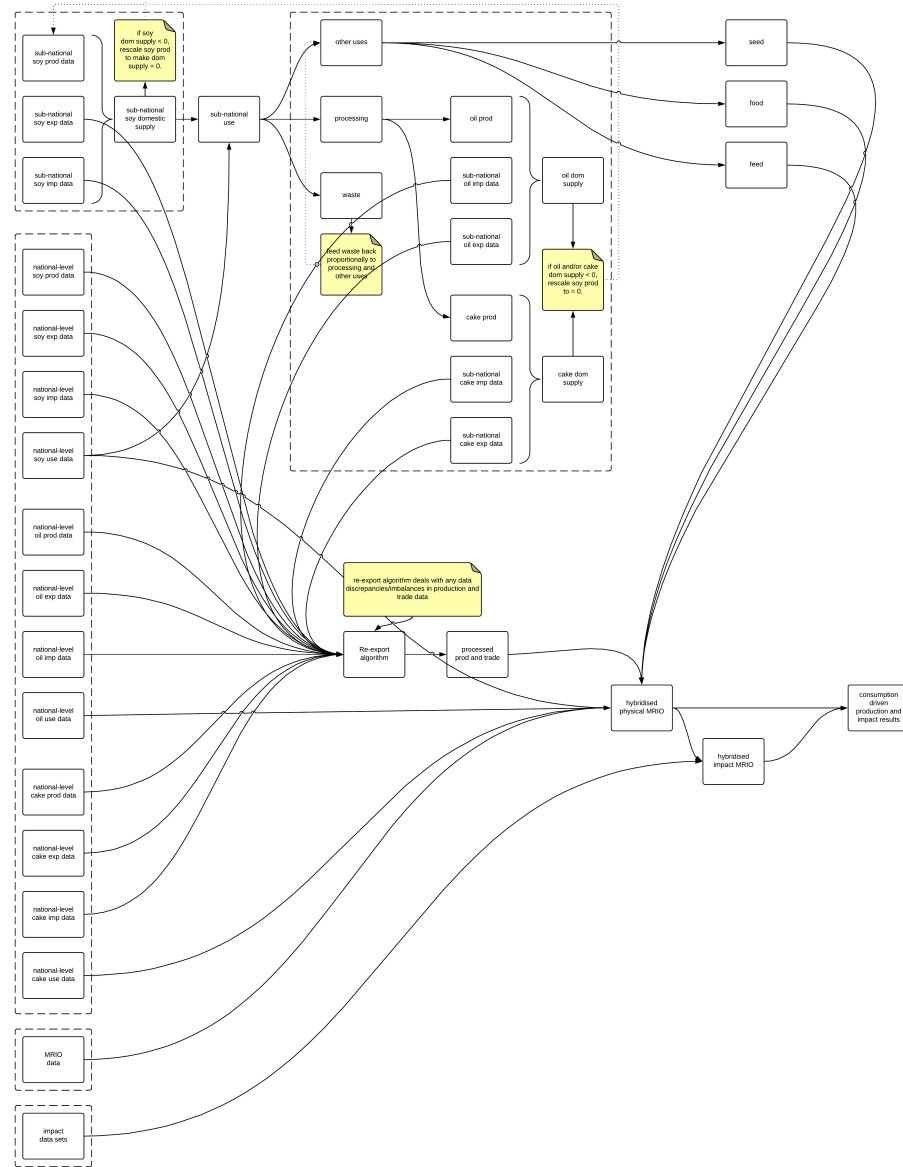
B   Soy model implementation

Figure 5: Visual representation and flow diagram of all data sources and interme-
diate steps in the application of IOTA to sub-national Brazilian soy production
within a global production and consumption context. The final outputs are
hybridised versions of the GTAP MRIO, linking consumption from economic
sectors across different sectors/regions to localised soy production/associated
land use.

26