AISHA: A tool to construct a series of contingency tables

Arne Geschke[a,*], Manfred Lenzen[a], Keiichiro Kanemoto[a] and Daniel Moran[a]

[a]*ISA, School of Physics A28, The University of Sydney NSW 2006, Australia*

Abstract

AISHA is a tool that facilitates building series of contingency tables. Users provide conventional monetary and/or environmentally extended datasets (for example input-output matrices) for regions to be considered in the contingency table. The software merges these datasets and formulates the matrix balancing constraints and optional, user-defined additional constraints as a constrained optimization problem. AISHA includes optimisation routines to solve the optimsation problem and exports the result for further use. This paper describes the methodolgy of AISHA including the preparation of an intial estimate, the formulation of the constraints matrix, boundary conditions, and the resulting optimisation problem. A graphical user interface (AISHA-GUI) is presented which permits the user to define all settings and to control the workflow of AISHA. Special focus is placed on the meta language that was developed to formulate the constraints. Finally, a small test example is presented including all necessary input paramters and results.

1 Introduction

1.1 Motivation and Aims

Why is there a need for a tool to construct contingency tables? The development of AISHA was mainly driven by a particular application from the discipline of Economics - input-output analysis - to an important contemporary problem - climate change.

Input-output (IO) analysis was conceived by Nobel Prize laureate Wassily Leontief around the second World War (Leontief 1936, Leontief 1941, and Leontief et al. 1953), and now forms a branch of economics, with its own international association[1] and designated journal[2]. Whilst initially thought for economic analysis, Leontief himself extended the technique to environmental applications (Leontief and Ford 1970, Leontief and Ford 1971), and since then it has been employed in thousands of empirical and theoretical studies (Hoekstra 2010).

Input-output theory is based on a particular type of contingency tables - input-output tables - that are now regularly published for more than 100 countries around the world. In addition, tailored sub-national or international multi-regional input-output (MRIO) tables are constructed by a number of academic teams.

Each statistical agency and research team requires some computational means of constructing IO tables. National agencies deal with relatively small tables and associated computer software

---

and hardware often exist in the form of desktop computer applications. World MRIO developers deal with much larger tables, and as a result there exists only a handful of initiatives around the world engaged with sophisticated yet non-standardised computational techniques (Wiedmann et al. 2011). These initiatives focus on providing a database for spatially-explicit economic applications'[3], Computable General Equilibirum modeling of global trade[4], analysis of environmental externalities[5], or implications of international trade for the spatial distribution of labour and skills[6].

The AISHA project is motivated by the increasingly important nexus of international trade and responsibility for greenhouse gas emissions[7]. During the past decade, our understanding of climate change has improved, but with it, the future outlook has worsened, in part due to newly discovered positive feedbacks (Luthje et al. 2006; Walter et al. 2006) and lower estimates for the absorptive capacity of the biosphere (Schuster and Watson 2006; Heimann and Reichstein 2008). At the same time, global emissions have during recent years approximated the more pessimistic emissions scenarios of the IPCC (Nakićenović and Swart 2000; van Vuuren and Riahi 2008). Compared to a decade ago, the problem of climate change is now perceived as more severe, more urgent, and as a result more political. The latter is reflected in increasing debates about the national responsibilities for the damages expected from climate change (Munksgaard and Pedersen 2001; Peters 2008; Peters and Hertwich 2008; Serrano and Dietzenbacher 2010). In particular, exporters of emissions-intensive commodities now argue more strongly than ever for a consumer-responsibility principle ("We produce products and these products are consumed by other countries, especially the developed countries. This share of emissions should be taken by the consumers but not the producers", Li Gao, Climate Change Negotiator for China, BBC News 2009). In response to these recent trends, various accounting, labelling, reporting, life-cycle, and policy frameworks for consumer responsibility have been created or revived (see Section 3 in Rueda-Cantuche et al. 2009), and some of these have dealt with international trade, such as the European EIPOT project (Wiedmann et al. 2009).

In order to underpin these initiatives, a comprehensive and reliable multi-region input-output (MRIO) database on emissions and international trade is necessary (Tukker et al. 2009, Wiedmann 2009). Such databases should ideally cover the entire world at high sector detail, so that emissions-intensive industries or commodities can be singled out. For example, previous multi-region studies have used either sector-disaggregated models for a limited number of countries, or sector-aggregated models for the world (Wiedmann et al. 2007; Moran et al. 2009; Wiedmann 2009). However, the MRIO initiatives listed above did not provide for maximum sector disaggregation, but instead opted for a breakdown into around 50 regions and 120 sectors common to all countries. Further, some initiatives did not attempt a continuous time series, only envisage one (basic price) valuation, and did not provide information on reliability and uncertainty (see Andrew et al. 2009, Appendix B). Finally, most initiatives required significant manual labour and time to complete one MRIO table.

The original aim of this project was therefore to go beyond existing ambitions for MRIO com-

---

[3]IDE-JETRO (`http://www.ide.go.jp/English/`)

[4]GTAP (`https://www.gtap.agecon.purdue.edu/`)

[5]EXIOPOL (`http://www.feem-project.net/exiopol/`)

[6]WIOD(`http://www.wiod.org/`)

[7]Andrew and Forgie (2008), Hertwich and Peters (2009), Lennox et al. (2010), Munksgaard and Pedersen (2001), Peters (2008), Peters and Hertwich (2008), Rodrigues and Domingos (2008), Tukker et al. (2009), Weber and Matthews (2007), Weber and Matthews (2007), Zhou (2010), Zhou (2009), Zhou et al. (2009).

pilation. Key goals were:

- **Detail**: Disaggregation of countries and sectors to the maximum possible level of detail, in order to assist environmental life-cycle and footprint-type assessments of international trade in the most accurate way possible;

- **Timeliness**: Continuous updating of the entire database, so that user analyses are relevant at the time;

- **Budget**: Implementation of the entire compilation and updating capability using less than 12 person-years initially, and less than 2 person-years per year continually;

- **Reliability**: Provision of data for constraint adherences in order to inform expert users and statisticians about the discrepancies between the fully balanced MRIO and disparate raw data;

- **Openness**: Public, free availability for research purposes, so that there is no barrier for wide dissemination;

- **Transparency**: Minimisation of assumptions made during the compilation (such as ratios of purchasers to basic prices), and close adherence to the raw data;

- **Uncertainty**: Provision of standard deviation estimates for all MRIO elements in order to aid comparative assessments, hypothesis testing, and decision-making;

- **Flexibility**: Compilation of table sheets expressed in basic prices as well as margins and taxes, and in current and constant US$, so that calculations for different purposes can be carried out;

- **Dynamics**: Creation of a historical time series back to 1990, in order to allow trend and scenario analyses, and projections.

The tool described in this report is currently used to create a time series with the following characteristics:

- **Detail**: Disaggregation into 160 countries at a detail of 25-500 sectors;

- **Timeliness**: Ability to continuously update the entire database with about 1 year delay;

- **Budget**: Development cost of 12 person-years and 50,000 AU$ of computer hardware and data;

- **Reliability**: Graphical User Interface controlling complete table construction runs, delivering tabular and visualised information on data quality and constraint adherence;

- **Openness**: Eora MRIO database freely accessible at `www.worldmrio.com` (from 1 January 2012);

- **Uncertainty**: Routine calculation of standard deviation estimates for all MRIO elements;

- **Flexibility**: Availability of table sheets in basic prices as well as 2 margins, taxes, and subsidies;

- **Dynamics**: Existence of a harmonised historical time series spanning 2000-2009.

Whilst the original aim of the AISHA project was solely to provide a tool for estimating large global MRIO tables including environmental accounts, it was subsequently recognised that a more general approach would also benefit applications using single-region IO tables, or even any kind of contingency tables. Hence, AISHA was developed to cope with contingency tables of any structure. An additional aim was to develop AISHA not only as a construction tool, but also as an information system that allows access to, and visualisation of raw data, construction parameters, diagnostics, and final results. Hence, in a sense, AISHA has espoused some of Leontief's early visions: an information system for the world economy (Leontief 1974; Leontief 1986a).

This document describes how the AISHA project has achieved its goals through a high level of procedural standardisation, automation, and data organisation. The following Section gives a brief but illustrative example of one class of contingency tables: Multi-Region Input-Output tables. Throughout this report, this example will be used to introduce the basic terminology, and for describing procedures and challenges in constructing time series of contingency tables (Section 2), and the functionality of AISHA (Section 3).

1.2   Multi-Region Input-Output accounts: An example for contingency tables

An Input-Output (IO) table shows the interdependencies between economic sectors, the government, the capital sectors, and households. For example it shows for a given national economy how much iron ore the steel industry buys in order to make steel, and how much steel the car industry buys in order to make cars. For an introduction into input-output theory see Leontief (1986b), Miller and Blair (1985), or Dixon (1996).

	Iron Ore	Steel	Cars	
Iron Ore		$		
Steel			$	
Cars				

Figure 1: Input-output table for a given economy.

An input-output table can be regionalised, that is the national economy can be broken down into sub-regions with the same industry sector breakdown (Figure 2). Such a Multi-Region Input-Output (MRIO) table shows for example how much NSW steel the Victorian car industry buys in order to make cars. (For an example of estimating an Australian sub-national MRIO see Gallego and Lenzen 2009).

Instead of, or in addition to showing sub-regions of one country, an input-output table can be extended to show a system of many regions (in this case: many countries, Figure 3), see Zhou (2010). Such an international MRIO system can in principle be set up for the entire world. It shows for example how much Australian iron ore German steel makers buy in order to make their steel, and how much Chinese steel Japanese car manufacturers buy in order to make their cars. The diagonal blocks of an MRIO table are the input-output tables of national economies, and the off-diagonal blocks are the trade matrices (Polenske 1980). For an introduction into

		NSW				VIC				QLD			
		Iron Ore	Steel	Cars		Iron Ore	Steel	Cars		Iron Ore	Steel	Cars	
NSW	Iron Ore		$				$				$		
	Steel			$				$				$	
	Cars												
VIC	Iron Ore		$				$				$		
	Steel			$				$				$	
	Cars												
QLD	Iron Ore		$				$				$		
	Steel			$				$				$	
	Cars												

Figure 2: Regionalised national inter-industry table.

MRIO theory see Leontief (1986b) and Miller and Blair (1985).

Input-output tables can be expressed in basic prices, or in producers' prices or purchasers' prices. The difference between these table variants is in their valuation. Basic prices are farm- or factory-gate prices of the respective commodity. Producers' prices are basic prices plus taxes on products minus subsidies on products. Purchasers' prices are producers' prices plus so-called margins, which are mark-ups on top of the basic price. Margins reflect the need to transport, insure, store, wholesale or retail commodities. Each transaction in a basic-price input-output

			AUS							DEU					JPN									RoW
			Ind			Com			FD	Ind		Com		FD	Ind				Com				FD	1
			1	2	3	1	2	3	1	1	2	1	2	1	1	2	3	4	1	2	3	4	1	1
AUS	Ind	1				x	x	x																
	Ind	2				x	x	x																
	Ind	3				x	x	x																
	Com	1	x	x	x				x	x	x			x					x	x	x	x	x	x
	Com	2	x	x	x				x	x	x			x					x	x	x	x	x	x
	Com	3	x	x	x				x	x	x			x					x	x	x	x	x	x
	VA	1	x	x	x																			
DEU	Ind	1	x	x	x				x	x	x			x					x	x	x	x	x	x
	Ind	2	x	x	x				x	x	x			x					x	x	x	x	x	x
	Com	1																						
	Com	2																						
	VA	1								x	x													
JPN	Ind	1																						
	Ind	2																						
	Ind	3																						
	Ind	4																						
	Com	1	x	x	x				x	x	x			x					x	x	x	x	x	x
	Com	2	x	x	x				x	x	x			x					x	x	x	x	x	x
	Com	3	x	x	x				x	x	x			x					x	x	x	x	x	x
	Com	4	x	x	x				x	x	x			x					x	x	x	x	x	x
	VA	1																	x	x	x	x		
RoW	1	1	x	x	x				x	x	x			x					x	x	x	x	x	x

Figure 3: An international MRIO table containing Australia, Germany, Japan, and the Rest of the World.
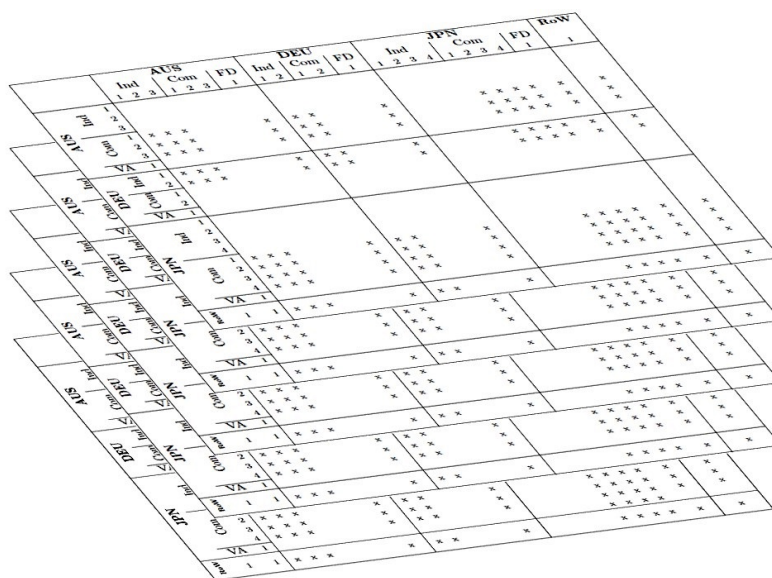
Figure 4: Stack of input-output tables representing basic prices, margins and taxes.

table has its corresponding margins and net taxes, and therefore an entire input-output table can be expressed in any of those valuation systems. In general, a complete set of tables comprises the basic-price table, and all margins and tax tables. They can be represented as a "stack" of tables (Figure 4; Dalgaard and Gysting 2004). An MRIO table can be generalised by adding rows underneath the table (Figure 5). These rows can contain information about employment, greenhouse gas emissions, land use, water use, or occupational health and safety incidents, and many other physical quantities (citealthoekstra2010). An entry in this additional (satellite) account block is for example the greenhouse gases emitted by the Chinese steel industry, or the employment in the US car manufacturing industry. Whilst the table parts in Figures 1-4 are expressed in monetary units, for example $, £, ¥, € etc, the satellite block is expressed in physical units such as person-years of employment, tonnes of greenhouse gas emissions, hectares of land, Megalitres of water, number of accidents, etc. The environment-economy link in IO analysis was first proposed by Leontief and Ford (1970). For an introductory description of this link see Duchin and Steenge (1999) and Forssell (1998).

Note that the satellite block in physical units exists only as an extension of the basic price sheet, and not of the mark-up (margins and tax) sheets (Figure 6). So-called generalised (that is physically extended) IO analysis has become an important tool in the disciplines of Industrial Ecology (Duchin 1992) and Life-Cycle Assessment (Lave et al. 1995, Heijungs and Suh 2002, Suh and Huppes 2005 and Suh and Nakamura 2007).

Finally, such generalised MRIO tables can be constructed for any year, and even arranged into a times series (Figure 7).
Input-output tables can come as symmetrical tables, distinguishing industry sectors, or they can be asymmetrical and show commodities in their rows and industries in their columns (so-called use tables), or vice versa (so-called supply tables). A system that combines both types into one is called a supply-use framework (see for example Gigantes 1970 and Madsen and Jensen-Butler 1999).

			AUS Ind			AUS Com			AUS FD	DEU Ind		DEU Com		DEU FD	JPN Ind				JPN Com				JPN FD	RoW
			1	2	3	1	2	3	1	1	2	1	2	1	1	2	3	4	1	2	3	4	1	1
AUS	Ind	1				x	x	x																
		2				x	x	x																
		3				x	x	x																
	Com	1	x	x	x				x	x	x			x					x	x	x	x	x	x
		2	x	x	x				x	x	x			x					x	x	x	x	x	x
		3	x	x	x				x	x	x			x					x	x	x	x	x	x
	VA	1	x	x	x																			
DEU	Ind	1	x	x	x				x	x	x			x					x	x	x	x	x	x
		2	x	x	x				x	x	x			x					x	x	x	x	x	x
	Com	1																						
		2																						
	VA	1								x	x													
JPN	Ind	1																						
		2																						
		3																						
		4																						
	Com	1	x	x	x				x	x	x			x					x	x	x	x	x	x
		2	x	x	x				x	x	x			x					x	x	x	x	x	x
		3	x	x	x				x	x	x			x					x	x	x	x	x	x
		4	x	x	x				x	x	x			x					x	x	x	x	x	x
	VA	1														x	x	x	x						
RoW	1	1	x	x	x				x	x	x			x					x	x	x	x	x	x
Employment			x								x									x				
Climate				x										x							x			
Water						x									x									x
Energy							x								x									x
Land Use								x											x				x	
...																								
...																								
...																								

Figure 5: Generalised MRIO table

The above features that apply to input-output tables also apply to Social Accounting Matrices (SAMs). The difference between an input-output table and a SAM is that the latter contains account entries that link the government, the capital sector and households with each other. Such entries are traditionally absent from input-output tables (Stone 1986; Roland-Holst 1990; Pyatt 1991; Hewings and Madden 1995).

The logical structure of this section suggests the following 8-dimensional hierarchy for addressing MRIO table entries:

- Time (year or financial year);

- Valuation (basic prices, producers' prices, purchasers' prices, margins, and taxes);

- Region of flow origin (country, state, region, etc);

- Entity of flow origin (industry, commodity, or physicals);

- Sector of flow origin (industries, commodities, or physical quantities);

- Region of flow destination (country, state, region, etc);

- Entity of flow destination (industry or commodity);

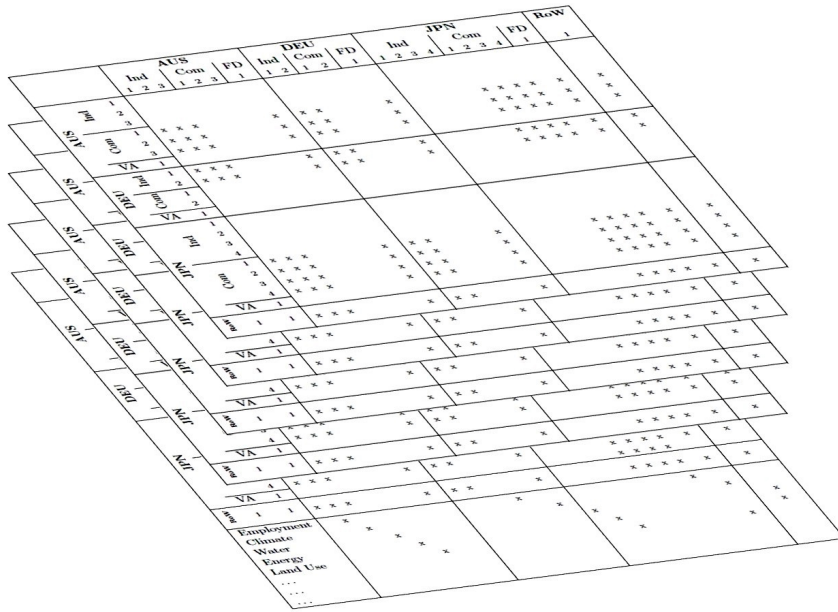- Sector of flow destination (industries or commodities).

Figure 6: Complete stack of SAM tables in different valuations, including a satellite block.
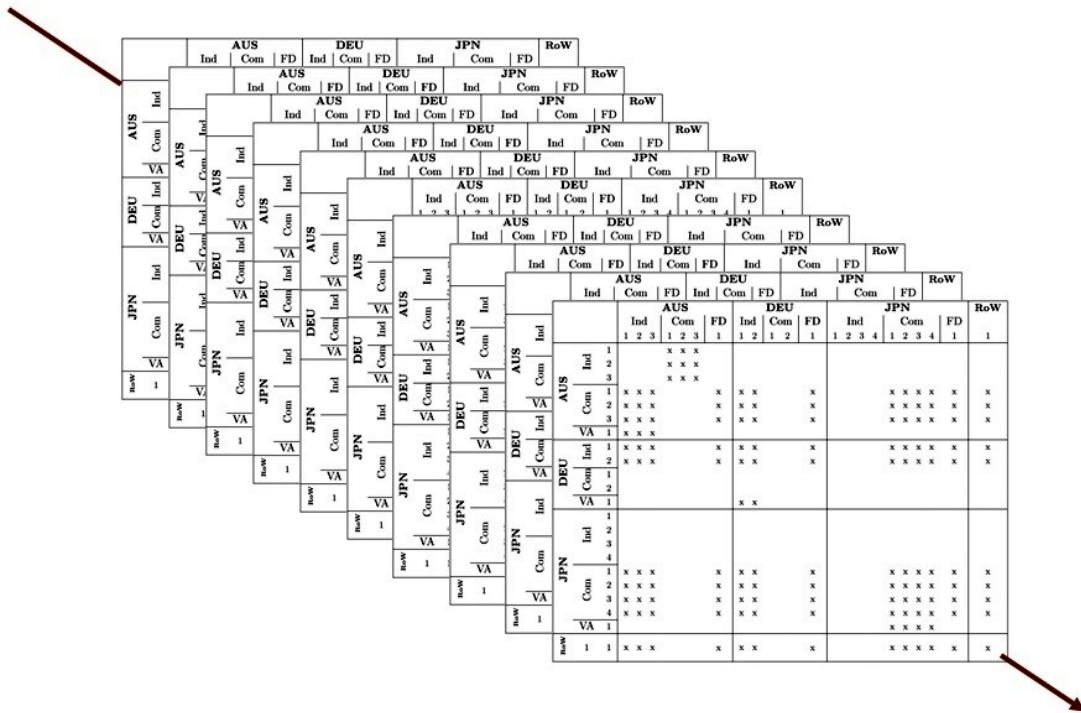

Figure 7: An MRIO time series.

This is just one example for what we will refer to as the *tree structure* of a contingency table (see Section 3.5).

2 Constructing a Series of Contingency Tables

2.1 Constrained Contingency Tables

The term "contingency" already reflects the existence of conditions to which such tables adhere. In fact, contingency tables are always subject to constraints posed by external data. This section reviews characteristics of constrained contingency tables.

Early applications of contingency table estimation concentrated on the problem of adjusting a table to known row and column sums, so-called marginal totals (Bacharach 1965, Deming and Stephan 1940, Friedlander 1961, Stephan 1942). This particular problem led Bacharach (1970) to develop the still widely used RAS method. The RAS method - in its basic form - bi-proportionally scales a matrix $\mathbf{T}^0$ of un-balanced preliminary estimates of an unknown real matrix $\mathbf{T}$, using $\mathbf{T}$'s known row and column sums. The balancing process is usually aborted when the discrepancy between the row and column sums of $\mathbf{T}^0$ and $\mathbf{T}$ is less than a previously fixed threshold. In essence, the known row and column totals represent contraints on the matrix to be estimated, and the adjustment to these constraints is an optimisation problem.

A slightly more complicated situation arises when some of the matrix elements of $\mathbf{T}$ are known in addition to its row and column sums, for example from an industry survey. The 'modified RAS' approach developed by Paelinck and Walbroeck (1963), Allen (1974), and Lecomber (1975) deals with such so-called partial information.

Moreover, situations can arise where, in addition to certain elements of $\mathbf{T}$, some aggregates of elements of $\mathbf{T}$ are known. For example, a published table $\mathbf{T}_{\mathrm{nat}}$ of national aggregates may constitute partial information when constructing a multi-regional input-output system, or a more disaggregated national table (Oosterhaven et al. 2008; Batten and Martellato 1985, p. 52-55; Gilchrist and St Louis 1999).

In the above approaches the partial information has to assume a particular aggregated form. A number of authors (Byron 1978, Morrison and Thuman 1980; van der Ploeg 1982; Cole 1992; Thissen and Löfgren 1998; Tarancon and Del Rio 2005) have generalised the method

1. by formulating it as a Lagrangian multiplier calculus, as an econometric estimation, or as an optimisation,

2. by enabling balancing of negative elements, and

3. by allowing for partial information of any form, by imposing constraints on arbitrary-sized and -shaped sets of elements of $\mathbf{T}^0$.

Lahr and de Mesnard (2004) provide an overview of methods, and Lenzen et al. (2006) present a comparison of constraint structures.

2.1.1 Structure of a Contingency Tables

In general, the subjects described by a contingency table will be classifiable according to some hierarchical system. The following definition provides an example for such a hierarchical classification for an MRIO.

Definition 2.1. *(8-DIM representation)*
In full generality, a multi-regional Input-Output Matrix (MRIO) can be addressed for example by using the following 8-dimensional structure:

1. Time (year or financial year), typically between 1 and 50 years;

2. Valuation (basic prices, producers' prices, purchasers' prices, margins, taxes, subsidies, free-on-board, cost-insurance-freight, etc), typically between 3 and 20 valuation types;

3. Region of flow origin (country, State, region, etc), typically between 1 and 250 regions;

4. Entity of flow origin (industry, commodity, or satellite), typically 1, 2 or 3 entities;

5. Sector of flow origin (industries, commodities, or satellite indicators), typically between 20 and 2,000 sectors;

6. Region of flow destination (country, State, region, etc), typically between 1 and 250 regions;

7. Entity of flow destination (industry or commodity), typically 1, 2 or 3 entities;

8. Sector of flow destination (industries or commodities), typically between 20 and 500 sectors.

We will refer to this 8-dimensional structure as the *8-DIM representation*. The position of each element within an MRIO can be uniquely addressed by an 8-tuple ξ_8 of the 8-DIM representation, defined as

$$\xi_8 = (x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8) \ .$$

In order for a series of MRIO tables to be harmonised, the region-entity-sector structure of 8-DIM dimensions 3 to 5, and 6 to 8, should neither change over time nor change across valuations. This requirement for harmonisation has proven extremely challenging for MRIO developers, mainly because the raw data underpinning MRIO construction is expressed in a multitude of different classifications. Assuming that international harmonisation of data sources cannot be expected in the short term, the AISHA tool was developed specifically with this challenge in mind.

In this report we define a standard MRIO table and use this table for exemplifying more general characteristics of contingency tables and their construction. In our standard MRIO the origin and destination structures are the industries, and commodities produced by those industries, in three particular countries: Australia, Germany, and Japan. For this example to be readable in this report, we aggregated the original sector classification of these countries into a sufficiently condensed form, representing each country by only a small number of sectors.

Example 2.2. (*Standard Example for MRIO*)
Australia is represented by a 3-sector supply-use table (SUT), Germany by a 2-sector industry-by-industry input-output table (IIOT), and Japan by a 4-sector commodity-by-commodity input-output table (CIOT). Finally, we add a 1-sector representation of the rest of the world (RoW).

Australia, Germany, and Japan were chosen because together these three countries cover all three possible input-output data structures (SUT, IIOT, and CIOT). Throughout this document we will use the following abbreviations in order to ensure the readability within tables.

Term	Abbreviation
Industries	Ind
Commodities	Com
Value Added	VA
Final Demand	FD

Special abbreviation rules apply to country names[8]. Australia's SUT-Table (Table 8), distinguishes 3 industry sectors, 3 commodity sectors and 1 final demand/value added sector.

Germany's IIOT-Table (Table 9), distinguishes 2 industry sectors, and 1 final demand / value

		Ind			Com		FD
Ind				x	x	x	
				x	x	x	
				x	x	x	
Com	x	x	x				x
	x	x	x				x
	x	x	x				x
VA	x	x	x				

Figure 8: Structure of the Australian SUT-Table. The 'x' indicate the positions of non-zero transaction values of the table.

added sector, but due to its IIOT-structure it has no commodity sectors.

Japan's CIOT-Table (Table 10), distinguishes no industry sectors (due to its CIOT structure),

		Ind		Com	FD
Ind	x	x			x
	x	x			x
Com					
VA		x	x		

Figure 9: Structure of the German IIOT-Table. The 'x' indicate the positions of non-zero transaction values of the table. Note that Germany has no commodity sectors.

but 4 commodity sectors, and 1 final demand/value added sector.

Additionally, the rest of the world (RoW) is identified as its own region, featuring only one entity and one sector.

Dimensions 3-5 and dimensions 6-8 of the 8-DIM representation are subject to the same internal hierarchy (Table 1). Note that hierarchical levels must not be countries, entities and sectors[9].

When combining several regions of origin/destination into an MRIO, the origin-destination hierarchies generate specific intra-country transaction blocks (origin = destination) and inter-country trade blocks (origin ≠ destination).

---

[8]Throughout this document country names are abbreviated using the official United Nation's 3-letter-country abbreviation. The offical country name abbreviations are available under `http://comtrade.un.org/db/mr/rfReportersList.aspx`.

[9]For example, in input-output satellite accounts expressed in physical units, these levels may be "CO_2-equivalent emissions", "gas types (CO_2, CH_4, N_2O, etc)" and "emission source (fuel combustion, fugitive, industrial processes, agriculture, land use changes)".

	Ind	Com				FD
Ind						
Com		x	x	x	x	x
		x	x	x	x	x
		x	x	x	x	x
		x	x	x	x	x
VA		x	x	x	x	

Figure 10: Structure of Japan's CIOT-Table. The 'x' indicate the positions of the non-zero transaction values of the table. Note that Japan has no industry sectors.

Region	Entity	Sectors
Australia (AUS) (1)	Industries (1)	Industry Sector 1: Primary, manufacturing and utilities
		Industry Sector 2: Trade and transport
		Industry Sector 3: other services
	Commodities (2)	Commodity Sector 1: Primary, manufacturing and utilities
		Commodity Sector 2: Trade and transport
		Commodity Sector 3: other services
	Value Added (3)	Value Added Sector 1
Germany (DEU) (2)	Industries (1)	Industry Sector 1: Primary, manufacturing and utilities
		Industry Sector 2: Services
	Commodites (2)	(No commodity sectors)
	Value Added (3)	Value Added Sector 1
Japan (JPN) (3)	Industries (1)	(no Industry sectors)
	Commodites (2)	Commodity Sector 1: Primary industries
		Commodity Sector 2: Manufacturing and utilities
		Commodity Sector 3: Trade and transport
		Commodity Sector 4: Other services
	Value Added (3)	Value Added Sector 1
Rest of the World (RoW) (4)	One Entity (1)	One Sector (1)

Table 1: Hierarchical structure of the *origin*-coordinates (*destination* coordinates are identical, except for final demand replacing value added) of an MRIO table, for our standard example including Australia, Germany, Japan, and Rest of the World.

Assume we want to address Germany's industry sector 2 (Services) as the sector of origin. In this case, the corresponding dimensions 3-5 in the 8-DIM representation are given by: region (Germany: 2), entity (Industry: 1), and Sector (Services: 2). Hence the corresponding 3-tuple for the dimensions 3-5 is $(2, 1, 2)$. The same approach can be taken for the destinations. Japan's commodity sector 1 (Primary industries), can be addressed by the 3 tuple $(3, 2, 1)$. Then, the dimensions 3-8 of the 8-DIM representation of a transaction between Germany's services sector and Japan's primary industries sector is given by the 6-tuple $(2, 1, 2, 3, 2, 1)$.

The 8-DIM representation of an MRIO transaction is completed by defining the year in which the transaction occurred (dimension 1) and the valuation sheet (dimension 2). Assume that dimension 2 (valuation) is set to 1 (basic-price sheet), then the 6-tuple of the dimensions 3-8 can be used to address any element in the basic price sheet of this example (Figure 3). Sheets representing valuations other than the basic price do not hold any value added data.

Graphically, the different valuations of an input-output table can be interpreted as a stack of sheets (Figure 4). For this example, there are 5 valuations (basic price, trade margin, transport margin, taxes on products, subsidies on products).

Finally, the first dimension is the years to which the MRIO refers. The 8-DIM representation allows multiple years to be considered at the same time (one stack as shown in Figure 4 per year). Assume that the first dimension is set to 1 (the first year counted from the base year), then the 8-tuple addressing the transaction value is

$$\xi_8 = (1, 1, 2, 1, 3, 2, 1) \ .$$

The corresponding transaction value within the contingency table $\mathbf{T}$ is referred to as t_{ξ_8}.

Finally, such generalised MRIO tables can be constructed for any year, and even arranged into a times series (Figure 11). However, usually only the MRIO for one year can be constructed at one time (see Tarancon and Del Rio (2005) for an attempt to constrain a series of input-output systems linked over time). The structure of such a single-year MRIO is shown in Figure 4.

2.1.2 Introducing Constraints on Contigency Tables

Contingency tables (and in particular IO-Tables) are often subject to constraints. A constraint is a mathematical requirement to elements of contigency tables. The requirements (or constraints) can be expressed in mathematical equations. Let

$$\mathbf{T} \ = \ \begin{pmatrix} t_{11} & t_{12} & t_{13} \\ t_{21} & t_{22} & t_{23} \\ t_{31} & t_{32} & t_{33} \end{pmatrix} \tag{1}$$

be a contingency table. $\mathbf{T}$ does not reflect the total structure given in example 2.2, but for the time being and for the sake of providing a simple example, we can interpret $\mathbf{T}$ as the basic price sheet of Germany's IIOT as it is shown in table 9. Following Table 1 the two industry sectors simply represent *Primary, manufacturing, and utilities (PriMaU)* and *Services*. Then $\mathbf{T}$ has the structure as shown in Table 2.

Perhaps the most important constraints on MRIOs are the so-called *balancing constraints*. In a balanced table, the sum over all elements of the i-th row must equal the sum over all elements of
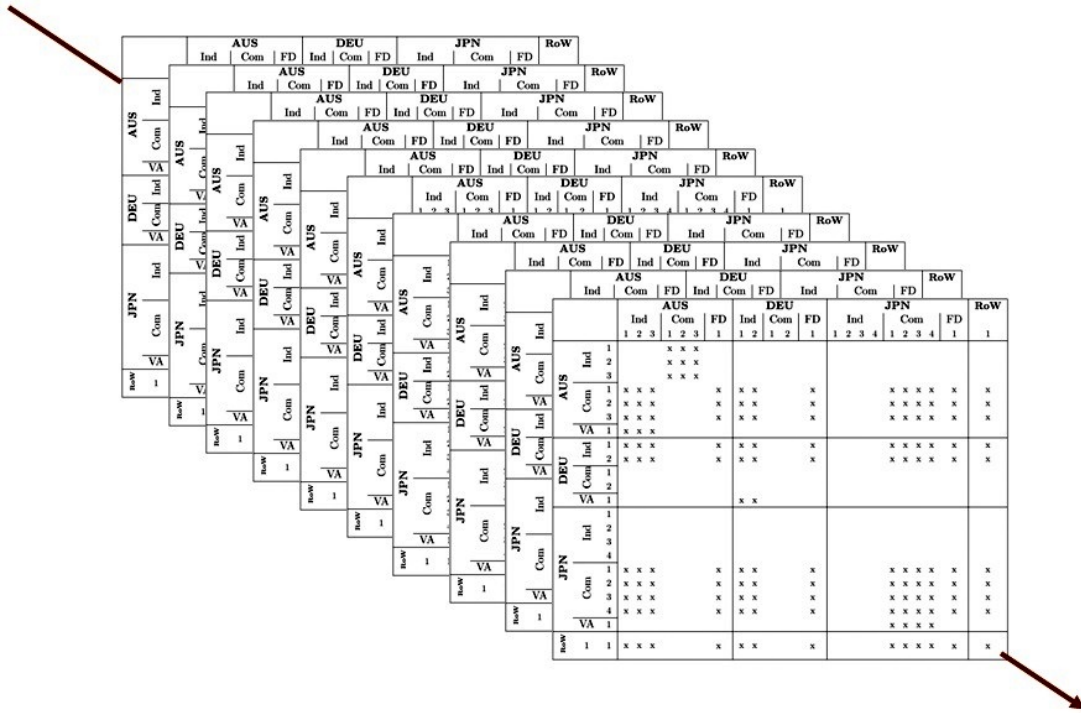
Figure 11: An MRIO time series.

	PriMaU	Services	Final Demand
PriMaU	t_{11}	t_{12}	t_{13}
Services	t_{21}	t_{22}	t_{23}
Value Added	t_{31}	t_{32}	

Table 2: **T** as Germany's IIOT-Table

the i-th column. For the contingency table **T** introduced in Equation 1 the balancing constraints are given (in principle, net taxes on products being included in value added) by a set of equations

$$t_{11} + t_{12} + t_{13} = t_{11} + t_{21} + t_{31}$$
$$t_{21} + t_{22} + t_{23} = t_{12} + t_{22} + t_{32}$$
$$t_{31} + t_{32} + t_{33} = t_{13} + t_{23} + t_{33} \quad ,$$

which is equivalent to

$$t_{11} + t_{12} + t_{13} - (t_{11} + t_{21} + t_{31}) = 0$$
$$t_{21} + t_{22} + t_{23} - (t_{12} + t_{22} + t_{32}) = 0$$
$$t_{31} + t_{32} + t_{33} - (t_{13} + t_{23} + t_{33}) = 0 \quad .$$

Since each of the diagonal elements t_{ii} for $i = 1, 2, 3$ appears twice with alternating signs in the corresponding equation, these equations simplify to

$$
\begin{aligned}
t_{12} + t_{13} - t_{21} - t_{31} &= 0 \\
t_{21} + t_{23} - t_{12} - t_{32} &= 0 \\
t_{31} + t_{32} - t_{13} - t_{23} &= 0 \;.
\end{aligned}
\tag{2}
$$

Equations 2 are the mathematical representation of the requirement that the sums of elements in the i-th column must equal the sum over the elements of the i-th row.

Similarly, other requirements can be expressed using the same notation. IO tables also have the requirement that the sum of each column in the margins matrices (valuations 2 and 3 in our example) must equal 0. The resulting constraint equation is

$$
t_{1i} + t_{2i} + t_{3i} = 0 \qquad \forall i
\tag{3}
$$

where i denotes the column (destination) index in the margins tables.

In balancing constraint the right-hand side of constraint equations (for example equations 2 and 3) are 0, but this does not need to be the case. In many early examples of IO table estimation, information was available on the gross output of sectors (Bacharach 1965, Deming and Stephan 1940, Friedlander 1961, Stephan 1942). Suppose the gross output for sector *Primary, manufacturing, and utilities* is given by a monetary value x_1, then the corresponding constraint equation (sum over all elements of sector *Primary, manufacturing, and utilities* equals x_1) is

$$
t_{11} + t_{12} + t_{13} = x_1
\tag{4}
$$

Another important class of constraints are *ratio constraints*, that consider known ratios amongst table elements. For example, Andrew et al. (2010) applied ratio constraints during the estimation of an MRIO table centered on New Zealand (R. Andrew, personal communication, April 2011). More specifically, he imposed that the production structure of sectors in a table $\mathbf{T}$ to be constructed should not deviate from the production structure in a known table $\mathbf{T}^*$. To formulate this mathematically, we write the production structure of an economy represented by a transaction table $\mathbf{T}$ and gross output vector diagonalised as $\hat{\mathbf{x}}$ as

$$
\mathbf{A} = \mathbf{T}^{-1}\hat{\mathbf{x}} \;.
\tag{5}
$$

The elements a_{ij} are interpreted as the amount of input that sector j required from sector i per unit of j's gross output. Equality of production structures inherent in $\mathbf{T}$ and $\mathbf{T}^*$ can be expressed as

$$
\begin{aligned}
a_{ij} \;&=\; \frac{t_{ij}}{\sum_i t_{ij} + \sum_l v_{lj}} = \frac{t_{ij}^*}{\sum_i t_{ij}^* + \sum_l v_{lj}^*} \;=\; : a_{ij}^* \\
\Longleftrightarrow \quad & t_{ij} - a_{ij}^* \sum_i t_{ij} - a_{ij}^* \sum_l v_{lj} = 0 \qquad \forall i, j
\end{aligned}
$$

Virtually every constraint that will be considered within this document is a combination of the basic constraints examples presented in this section. Combining these basic structures can result in constraints of almost arbitrary complexity, especially for large contigency tables. Examples

can be found throughout this document.

Finally, contingency tables may also be subject to boundary constraints. Unlike the constraints discussed so far (which were all equality constraints), boundary constraints are inequality constraints that define a lower or upper bound for elements of the contingency tables. The most common boundary constraints in MRIO tables are restrictions on certain parts of the different valuation sheets. Basic-price intermediate demand for example should only have positive values. A subsidies sheets (our example's sheet 5) must only have negative values. Finally, changes in inventories are a component of final demand that do not have any boundary restrictions, because sectors in an economy can in subsequent years either add to, or draw from stocks, which results in positive or negative changes in inventories.

Hence for each element t_{ij} of $\mathbf{T}$ there may be numbers l and u that define bounds for t_{ij}, i.e.

$$l \leq t_{ij} \leq u \tag{6}$$

2.1.3 Vectorisation of the 8-DIM representation

Throughout this section, let

$$\xi_8 = (x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8) \in \mathbb{N}^8$$

be the 8-DIM representation of an arbitrary element of a contingency table $\mathbf{T}$, let t_{ξ_8} be the transaction value of $\mathbf{T}$ addressed by the 8-tuple ξ_8, let I be the total number of elements within $\mathbf{T}$, and let

$$\underline{I} := [1, I] \cap \mathbb{N} = \{1, \ldots, I\}$$

be the set that contains all natural numbers from 1 to I.

In this section we are developing a vectorised representation of $\mathbf{T}$. Let $D \in \mathbb{N}^8$ be the finite subset of 8-tuples that are required to represent the contingency table by the 8-DIM representation. Then we are looking of a bijection ψ that fullfils

$$\psi : D \to \underline{I}, \qquad \psi \text{ is bijective.} \tag{7}$$

In order to achieve this goal, we will firstly introduce a more intuitive representation for each valuation sheet of a contingency table. Figure 3 shows the basic price sheet of Example 2.2 which is represented by the last 6 dimensions (elements $x_3, \ldots, x_8$ of ξ_8) of the 8-DIM representation. Our first aim is to find a 2-dimensional representation of each valuation sheet.

Definition 2.3. *(4-DIM representation of a contigency table)*
Let (x_3, x_4, x_5) be the 3-tuple of dimension 3, 4, and 8 of the 8-DIM representation. Then (x_3, x_4, x_5) refers to the region of origin, entitiy of origin and sector of origin of a transaction value within the contingency table. Let us define

- for every feasible x_3, let E_{x_3} be the total number of entities for region x_3, and

- for every feasible x_4, let S_{x_3, x_4} be the total number of sectors within the x_4-th entity of region addressed by x_3.

Then we can define

$$\phi_1 : \ \mathbb{N}^3 \to \mathbb{N} : \ (x_3, x_4, x_5) \mapsto \sum_{i_1=1}^{x_1-1} \sum_{i_2=1}^{E_{i_1}} S_{i_1,i_2} + \sum_{i_3=0}^{x_4-1} S_{x_3,i_3} + x_5 \tag{8}$$

Equation 8 maps dimensions 3,4, and 5 of the 8-DIM representation (the three origin dimensions) bijectively onto the set $\underline{K}$, where K is the total amount of origin sectors of one valuation sheet of the contingency table.

The concept be be applied to (x_6, x_7, x_8) (dimensions 6,7, and 8 of the 8-DIM representation, the destination dimensions), yielding a function

$$\phi_2 : \ \mathbb{N}^3 \to \mathbb{N} \tag{9}$$

For ξ_8, we can then define:

$$\gamma : \ \mathbb{N}^8 \to \mathbb{N}^4 : \ (x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8) \mapsto \big(x_1, x_2, \phi_1(x_3, x_4, x_5), \phi_2(x_6, x_7, x_8) \big) \tag{10}$$

With $k := \phi_1(x_3, x_4, x_5)$ and $l := \phi_2(x_6, x_7, x_8)$ we can re-write ξ_8 as

$$\xi_4 := (x_1, x_2, k, l) \in \mathbb{N}^4 \ .$$

The corresponding transaction value within the contingency table $\mathbf{T}$ is referred to as t_{ξ_4}.

This reduces the original 8-tuple of the 8-DIM representation to a 4-tuple. This representation is called the *4-DIM representation*.

Our next step will be to reduce the 4-DIM representation to a 3-dimensional representation.

Definition 2.4. *(Vectorised representation of a valuation sheet of a contigency table)*
Let $\xi_4 = (x_1, x_2, k, l)$ be a 4-DIM representation of an arbitraty element of a contingency table $\mathbf{A}$. The for each fixed pair (x_1, x_2) the tuple (k, l) represents a valuation sheets of the contingency table.

Let $\mathbf{V}$ be a valuation sheet of a contingency $\mathbf{T}$ table of the form

$$\mathbf{V} \ = \ \begin{pmatrix} v_{11} & v_{12} & \dots & v_{1L} \\ v_{21} & v_{22} & \dots & v_{2L} \\ \vdots & \vdots & \ddots & \vdots \\ v_{K1} & v_{K2} & \dots & v_{KL} \end{pmatrix} \tag{11}$$

and let v_{kl}, $k = 1, 2, \dots, K$, $l = 1, 2, \dots, L$ denote the elements in $\mathbf{V}$. Then we define the *vectorisation* of $\mathbf{V}$ as

$$\phi_3 : \mathbb{N}^2 \to \mathbb{N} : \ (k, l) \mapsto (k - 1) \cdot L + l \ . \tag{12}$$

With Equation 12 the *vectorisation* ν of $\mathbf{V}$ is

$$\nu \in \mathbb{R}^{(K \cdot L)}, \nu_{\phi_3(k,l)} := v_{kl} \qquad \Longrightarrow \qquad \nu = \begin{pmatrix} v_{11} \\ v_{12} \\ \vdots \\ v_{1L} \\ v_{21} \\ v_{22} \\ \vdots \\ v_{2L} \\ \vdots \\ v_{k1} \\ v_{k2} \\ \vdots \\ v_{KL} \end{pmatrix} \tag{13}$$

Using function ϕ_3, we can reduce the 4-DIM representation to a 3-dimensional representation of the contingency table by defining

$$\beta : \mathbb{N}^4 \to \mathbb{N}^3 : \ (x_1, x_2, k, l) \mapsto \big(x_1, x_2, \phi_3(k, l) \big) \ . \tag{14}$$

Function ϕ_3 is fact a bijection, and therefore β is a bijection.

By defining $j := \phi_3(k, l)$ we obtain

$$\xi_3 := (x_1, x_2, j) \in \mathbb{N}^3 \ .$$

which is a 3-dimensional representation of the contingency table $\mathbf{T}$. The corresponding transaction value within the contingency table $\mathbf{T}$ is referred to as t_{ξ_3}.

Definition 2.5. *(1-DIM representation)*
The 3-dimensional representation of a contingency table $\mathbf{T}$ introduced in Definition 2.4 represents each valuation sheet as a vector. The sizes of the valuation sheets (and therefore the lengths of the vectorised representations) are equal for each feasible pair (x_1, x_2).

Suppose X_2 is the total amount of valuations per year within the contingency table, and ν is a vectorised representation of a valuation sheets with elements ν_j, $j = 1, \ldots, J$. Let (x_1, x_2, j) be the 3-dimensional representation of the contingency table $\mathbf{T}$. Then

$$\alpha : \mathbb{N}^3 \to \mathbb{N} : \ (x_1, x_2, j) \mapsto (x_1 - 1) \cdot X_2 + (x_2 - 1) \cdot J + j \tag{15}$$

maps the 3-tuple (x_1, x_2, j) bijectively to the set $\underline{I}$, where I is the total amount of elements in $\mathbf{T}$. By defining

$$i := \alpha\big((x_1, x_2, j)\big) = (x_1 - 1) \cdot X_2 + (x_2 - 1) \cdot J + j \tag{16}$$

we obtain the *1-DIM representation.*

Corollary 2.6. *(Complete Vectorisation of a Contingency Table)*
Let $\mathbf{T}$ be a contingency table and let

$$D = \{ \ \xi_8 \mid \xi_8 \text{ is an 8-DIM representation of a transaction value of } \mathbf{T} \} \subset \mathbb{N}^8$$

be the set of 8-tuples of the 8-DIM representation of $\mathbf{T}$. Let I be the total amount of elements in $\mathbf{T}$, then a bijective function for Equation 7 is given by

$$\psi : D \to \underline{I}; \ \psi := \alpha \circ \beta \circ \gamma \tag{17}$$

with functions α as defined in Equation 15, β as defined in Equation 14, and γ as defined in Equation 10.

Definition 2.7. *(Vectorisation of a Contingency Table)*
Using function ψ from Corollary 2.6 we obtain a vectorisation of the contingency table $\mathbf{T}$. Then we define

$$\tau \in \mathbb{R}^I : \ \tau_{\psi(\xi_8)} = t_{\xi_8} \tag{18}$$

as the *vectorisation* of the contingency table $\mathbf{T}$ with t_{ξ_8} being the elements of $\mathbf{T}$ addressed by corresponding ξ_8 in the 8-DIM representation.

Throughout this work, we will refer to the vectorisation of a contingency table $\mathbf{T}$ as τ and to the elements of τ as τ_i with $i = 1, \ldots, I$.

2.1.4 Formulating Constraints as Matrix Equations

All concepts for constraints presented in section 2.1.2 yield constraint equations that are *linear* in the coefficients of the contigency table. Assume that we have a contingency table $\mathbf{T}$ and its vectorisation τ, then we can express *every* constraint in the form

$$g^T \tau = c \qquad \text{with} \qquad g, \tau \in \mathbb{R}^I, \ c \in \mathbb{R} \tag{19}$$

The vector g contains the appropriate coefficients and the value r is the corresponding right hand side value of the constraint. Assume that the vectorised contigency table τ is subject to K constraints. Then each constraint can be represented in the form of Equation 19 with appropriate vectors g_k and right-hand-side values c_k, $k = 1, \ldots, K$. More importantly, all K constraints can be summarised as

$$\underbrace{\begin{pmatrix} g_1^T \\ g_2^T \\ \vdots \\ g_K^T \end{pmatrix}}_{=:\mathbf{G}} \tau = \underbrace{\begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_K \end{pmatrix}}_{=:c}$$

Hence all constraints can be elegantly written as

$$\mathbf{G}\tau = c \qquad \text{with} \qquad \mathbf{G} \in \mathbb{R}^{K \times I}, \ \tau \in \mathbb{R}^I, \ c \in \mathbb{R}^K \tag{20}$$

Definition 2.8. *(Active constraints)*
Let $\tau \in \mathbb{R}^I$ be a vectorised contingency table, $g \in \mathbb{R}^I$, and $c \in \mathbb{R}$. Let

$$g^T \tau = c$$

be a constraint as introduced in Equation 19. Let τ_i and g_i be elements of τ and g. The we call the constraint *active for* τ_i if $g_i \neq 0$. Hence, τ_i "is affected" by the constraint. We may alternatively say that the *constraint vector g is active on τ_i.*

Definition 2.9. *(Classification of Constraints Matrices)*
Let τ be the vectorised representation of a contingency table, let $\mathbf{G}$ be a constraints matrix for τ as given in Equation 20, and let g_j^T, g_k^T be rows of $\mathbf{T}$.

- We call $\mathbf{G}$ *non-degenerate*, if for each element τ_i of τ there is at most one constraint vector g_j^T which is active on τ_i.

- We call $\mathbf{G}$ *degenerate*, if for at least one element τ_i of τ there are at least two constraints vectors g_j^T and g_k^T which are active on τ_i and

$$g_j^T \neq g_k^T$$

holds.

As mentioned before, boundary constraints may hold for each element of the vectorised MRIO τ (as shown in Equation 6). Hence for each elements τ_i of τ there might be a lower bound l_i and an upper bound u_i. These boundary conditions can be expressed as

$$l_i \leq \tau_i \leq u_i$$

The values l_i and u_i can be summarised in vectors l and u, so that the boundary conditions for the whole MRIO τ is given by

$$l \leq \tau \leq u \qquad \tau, l, u \in \mathbb{R}_\infty^I \qquad \text{with} \qquad \mathbb{R}_\infty = \mathbb{R} \cup \{\infty, -\infty\} \tag{21}$$

By using $\mathbb{R}_\infty$ it is possible to include those values τ_i into Equation (21) that are subject to only one bound, or no bounds at all.

Example 2.10. *(Balancing Constraints for a 3×3-Matrix)*
Let us recall the contingency table $\mathbf{T}$ from Equation 1. The balancing constraints for $\mathbf{T}$ are given by Equations (2). We will vectorise $\mathbf{T}$ and express the balancing constraints as $G\tau = c$.

Using Equation 13, $\mathbf{T}$ becomes

$$\mathbf{T} \in \mathbb{R}^{3\times 3} \quad \implies \quad \tau \in \mathbb{R}^9, \ \tau = \begin{pmatrix} t_{11} \\ t_{12} \\ t_{13} \\ t_{21} \\ t_{22} \\ t_{23} \\ t_{31} \\ t_{32} \\ t_{33} \end{pmatrix}$$

The first balancing constraint from (2)

$$t_{12} + t_{13} - t_{21} - t_{31} = 0$$

can be rewritten as

$$
\underbrace{\begin{pmatrix} 0 & 1 & 1 & -1 & 0 & 0 & -1 & 0 & 0 \end{pmatrix}}_{=:g_1^T}
\begin{pmatrix} t_{11} \\ t_{12} \\ t_{13} \\ t_{21} \\ t_{22} \\ t_{23} \\ t_{31} \\ t_{32} \\ t_{33} \end{pmatrix} = 0
$$

The same strategy can be applied to the second and third equation of Equation 2, yielding g_2^T and g_3^T. Combining the three vectors g_1^T, g_2^T and g_3^T in a matrix form gives Equation 2 the structure

$$
\underbrace{\begin{pmatrix} 0 & 1 & 1 & -1 & 0 & 0 & -1 & 0 & 0 \\ 0 & -1 & 0 & 1 & 0 & 1 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 & 0 & -1 & 1 & 1 & 0 \end{pmatrix}}_{=\mathbf{G}}
\underbrace{\begin{pmatrix} t_{11} \\ t_{12} \\ t_{13} \\ t_{21} \\ t_{22} \\ t_{23} \\ t_{31} \\ t_{32} \\ t_{33} \end{pmatrix}}_{=\tau} = \underbrace{\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}}_{=c}
$$

Hence, the contigency table $\mathbf{T}$ and the balancing constraints given by Equations 2 can be represented in the form $\mathbf{G}\tau = c$.

Example 2.11. *(Source data for right hand side values)*
Recall Equation 4

$$
t_{11} + t_{12} + t_{13} = x_1
$$

for the same contingency table $\mathbf{T}$ as in Example 2.10. Assume that there is a data source for x_1. In order to consider the data for x_1, have to formulate Equation 4 a constraint equation. Using the same vectorisation for $\mathbf{T}$ as in Example 2.10 we obtain

$$
\begin{pmatrix} 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}
\begin{pmatrix} t_{11} \\ t_{12} \\ t_{13} \\ t_{21} \\ t_{22} \\ t_{23} \\ t_{31} \\ t_{32} \\ t_{33} \end{pmatrix} = x_1 \quad .
$$

2.1.5 Data Reliability

Data published by national[10] or international[11] agencies are associated with a number of error sources (Murray 1998, Quandt 1958, Smith et al. 1998), simply because every data item is based on some observation of a quantity. Every observation is characterised by a certain degree of measurement reliability, which can then be translated into an estimate of the uncertainty of the respective data item. The reason is that the data sources for instance even one single entry of a published national IO-Table is based on the evaluation of possibly several hundreds of values. Assuming that observations of a quantity are normally or log-normally distributed around a mean, the stochastic uncertainty of published data, can be conveniently expressed using *standard deviation* σ for each published data point.

Uncertainty analysis has a long-standing tradition in input-output analysis[12], especially with view of elucidating the properties of stochastic input-output systems[13]. Lenzen (2000) and Wiedmann et al. (2010) present estimates of standard deviations for data items supporting the construction of input-output tables for Australia and the United Kingdom. A number of authors[14] describe strategies for estimating standard deviations of quantities that are derived from input-output tables. Others[15] apply input-output table uncertainty to Life-Cycle Assessment.

Uncertainty data are usually not available from sources such as national statistical agencies. Providing information about uncertainty is essential for researchers in judging the reliability of results they may have obtained by using contingency tables. Lenzen et al. (2010b) developed a RAS variant for determining standard deviations of (MR)IO tables, using an MRIO for the UK as a case study. In this RAS variant, they employed published data on the standard deviations of the right-hand-side vector c in order to infer standard deviations of the vectorised MRIO τ. By formulating an underdetermined optimisation problem, these authors solved for the unknown standard deviations of the transaction values of the UK MRIO.

Revisiting the vectorised representation of an IO table presented in Definition 2.4, we can formulate

Definition 2.12. (*Vector of standard deviations for a vectorised IO-Table*)
Let τ be a vectorised IO table as defined in Definition 2.4, and let σ_i be the corresponding standard deviation of τ_i. Then we call

$$\sigma_\tau = \begin{pmatrix} \sigma_1 \\ \sigma_2 \\ \vdots \\ \sigma_I \end{pmatrix}$$

the *vector of standard deviations* of τ.

---

[10] For example the *Australian Bureau of Statistics (ABS)* (`http://www.abs.gov.au/`) for Australia or *Bureau of Economic Analysis (BEA)* (`https://www.bea.gov/index.htm`) for the US.

[11] For example *United Nations Commodity Trade Statistics Database (UN Comtrade)* (`http://comtrade.un.org/`).

[12] Evans (1954), Gibbons et al. (1982), Murray (1998), Park et al. (1981), Park (1973), Quandt (1958), Stevens and Trainer (1980), Sonis and Hewings (1989).

[13] Goicoechea and Hansen (1978), Hanseman (1982), Hanseman and Gustafson (1981), Kop Jansen (1994), Ten Raa and Rueda-Cantuche (2007), Ten Raa and Steel (1994), West (1986).

[14] Bullard and Sebald (1977); Bullard and Sebald (1988); Sonis and Hewings (1989); Sonis and Hewings (1995); Sakai et al. (2000); Peters (2007); Weber (2008); Lenzen et al. (2010b)

[15] Nansai et al. (2001); Yoshida et al. (2001); Hondo et al. (2002); Williams et al. (2009)

The same principle applies to constraints. An example is the gross output for sector *Primary, manufacturing, and utilities* (Equation 4) of a country which is represented by sum over a certain set of elements of the contingency table. Revisiting Equation 4

$$t_{11} + t_{12} + t_{13} = x_1 \qquad \text{with} \qquad x_1 \in \mathbb{R}$$

then the value x_1 is also subject to a certain reliability (for example: the GDP of a country cannot be measured or calculated with total accuracy), which is represented by the standard deviation σ_{x_1} for x_1.

For a larger number of constraints we can recall equation (20) and express the constraints as

$$\mathbf{G}\tau = c$$

The standard deviations for the vector of constraint values c can then be summarised in a vector σ_c of equal dimensions as c (similar to Definition 2.12).

Definition 2.13. *(Soft and hard constraints)*
Let τ be a vectorised contingency table and let $\mathbf{G}\tau = c$ be a set of constraints on τ and let $\sigma_c = (\sigma_{c,1}, \sigma_{c,2}, \ldots, \sigma_{c,K})^T$ be the vector of standard deviations for the vector of constraints values c. Then we categorise each row g_i^T of $\mathbf{G}$ as follows:

We say that g_i^T is a *hard constraint* if $\sigma_{c,i} = 0$. We call g_i^T a *soft constraint* if $\sigma_{c,i} \neq 0$.

2.1.6 The Reconciliation Problem

When designing a contingency table (for example an MRIO), the available data is usually published by different data providers. The individual data sets have to be assembled as one contingency table $\mathbf{T}^0$ from all these sources. In Example 2.2 the national data for Australia, Germany and Japan were supplied by the individual national agencies[16]. Figure 3 shows the structure of the basic price sheet of Example 2.2 which was assembled using data sets from these different sources.

The initial contingency table $\mathbf{T}^0$ is usually subject to external constraints. In Example 2.2, the initial MRIO is subject to balancing constraints (see Example 2.10). Suppose all constraints are represeneted by the constraints matrix $\mathbf{G}$. Then (following Definition 2.9) two of the following cases will hold:

1. $\mathbf{G}$ is non-degenerate, or

2. $\mathbf{G}$ is degenerate.

A reconciliation problem occurs for the second case ($\mathbf{G}$ is degenerate), because there is at least one element within the contingency table that is subject to more than one constraint. In this case we have to use an optimisation algorithm to obtain a contingency table $\mathbf{T}$ which does fulfill all constraints given by $\mathbf{G}$. Usually there are more elements in the contingency table than there are constraints. Hence, the unknowns outnumber the external constraints, resulting in the system being underdetermined, that is exhibiting too many degrees of freedom to be solved analytically.

---

[16]The data sources for Example 2.2 are as follows. Australia: Australian Bureau of Statistics, `http://www.abs.gov.au`; Germany: Statistisches Bundesamt Deutschland, `http://www.destatis.de`; Japan: Ministry of Internal Affairs and Communications, `http://www.soumu.go.jp/english/index.html`.

In order to approach this problem, one has to make assumptions about the unknown elements to provide initial data for each value in the MRIO, yielding a generally unbalanced initial estimate $\mathbf{T}^0$ for the contingency table[17].

Using $\mathbf{T}^0$, an optimisation problem to find a contingency table $\mathbf{T}$ that fulfills all constraints (a so-called feasible contingency table) can be formulated as

$$\min_{\tau} \; f(\tau, \tau^0) \qquad \text{subject to} \qquad \mathbf{G}\tau = c \; ,$$

where τ and τ^0 are the vectorised representations of $\mathbf{T}^0$ and $\mathbf{T}$. Often the problem is augmented by introducing vectors of upper and lower bounds l and u for τ.

$$\min_{\tau} \; f(\tau, \tau^0) \qquad \text{subject to} \qquad \mathbf{G}\tau = c, \qquad l \le \tau \le u \; . \tag{22}$$

Equation 22 represents the mathematical problem to find a feasible contingency table $\mathbf{T}$.

Note that Equation 22 does not specify the *objective function f*. In fact, the choice of a suitable and meaningful objective function *and* a powerful optimisation method that can solve Equation 22 has been a major topic in research for the last decades[18].

Regardless of which objective function is chosen for Equation 22, the following properties would be advantageous for the optimisation method used to solve the problem:

a) incorporate constraints on arbitrarily sized and shaped subsets of matrix elements, instead of only fixing row and column sums;

b) allow considering the reliability of the initial estimate;

c) allow considering the reliability of external constraints;

d) be able to handle negative values;

e) be able to handle conflicting external data.

Remark 2.14.
Equation 22 does not take into consideration the reliability of data as discussed in section 2.1.5. Some approaches do in fact consider the reliability of contingency table elements and constraint data. Hence, the vectors σ_τ and σ_c are taken into account as well. In this case, Equation 22 formally becomes

$$\min_{\tau} \; f(\tau, \tau^0, \sigma_\tau, \sigma_c) \qquad \text{subject to} \qquad G\tau = c, \qquad l \le \tau \le u \tag{23}$$

Since Equation 23 provides the most general approach to the problem of obtaining a contingency table, we will refer to this equation throughout this report. Equation 22 is essentially a special case of the more general Equation 23.

Remark 2.15. *(Passing Results on in Time Series Applications)*
For time series applications, the result of the the optimisation problem (Equation 23) is passed on to the subsequent time step as the initial estimate. Suppose for time step 1 of the time

---

[17]Bouwmeester and Oosterhaven (2008), Oosterhaven et al. (2008), Oosterhaven et al. (2011).
[18]Bacharach (1970) presents the method which has been the most successful one so far: the RAS method.

series Equation 23 delivers the result $\tau^{(1)}$. By passing $\tau^{(1)}$ on to the next time step 2, the corresponding optimisation problem for that time step becomes

$$\min \ f(\tau^{(2)}, \tau^{(1)}, \sigma_{\tau^{(2)}}, \sigma_c) \qquad \text{subject to} \qquad \mathbf{G}\tau^{(2)} = c, \qquad l \leq \tau^{(2)} \leq u \ .$$

Hence, for time series application, the user has to define a full initial estimate only for the base year. For each subsequent year the result of the previous year is used as the initial estimate.

2.2 Updating a Series of Contingency Tables

Especially large application obtain their input data from a large number of data sources. While some of the source data may be available at one point in time, other source data may be available at a later point. Hence, the data situation for a contingency table can vary over time.

Equation 23 can be applied at any given stage of data availability. The solution of the optimisation problem given by Equation 23 depends on the available input data. Hence, if new input data become available, the variables for Equation 23 change, and the solution for the equation has to be re-calculated with respect to the new variables in order to reflect the newly available input data in the result.

In time series applications, the optimisaton problem given by Equation 23 has to be solved for every time step in the time series. According to Remark 2.15, the result (say $\tau^{(1)}$) of Equation 23 for a particular time step (say t_1) is used as the initial estimate for Equation 23 at the subsequent time step (say t_2). If $\tau^{(1)}$ has to be re-calculated due to newly available input data, we obtain a new result for Equation 23 for time step t_1, say $\tau^{(1)}_{\text{new}}$. The result of Equation 23 for the time step t_2 does not reflect these new data, since the initial estimate for Equation 23 at time step t_2 was the original $\tau^{(1)}$. Therefore, the solution for Equation 23 at time step t_2 also has to be re-calculated, using $\tau^{(1)}_{\text{new}}$ as the initial estimate. Consequently, the results for each subsequent time step of the time series have to be re-calculated in order to consider the newly available input data for time step t_1. Especially for large systems, this can be quite time-consuming.

3   AISHA: A Tool to Construct a Series of Large Contingency Tables

3.1   Matlab-Based Graphical User Interface (GUI)

AISHA[19] is essentially a software package that is designed to perform the methodology to construct a series on contingency tables as it was described in Section 2. The functioning of AISHA is controlled by the user through a graphical user interface (GUI), referred to as the AISHA-GUI (Figure 12).
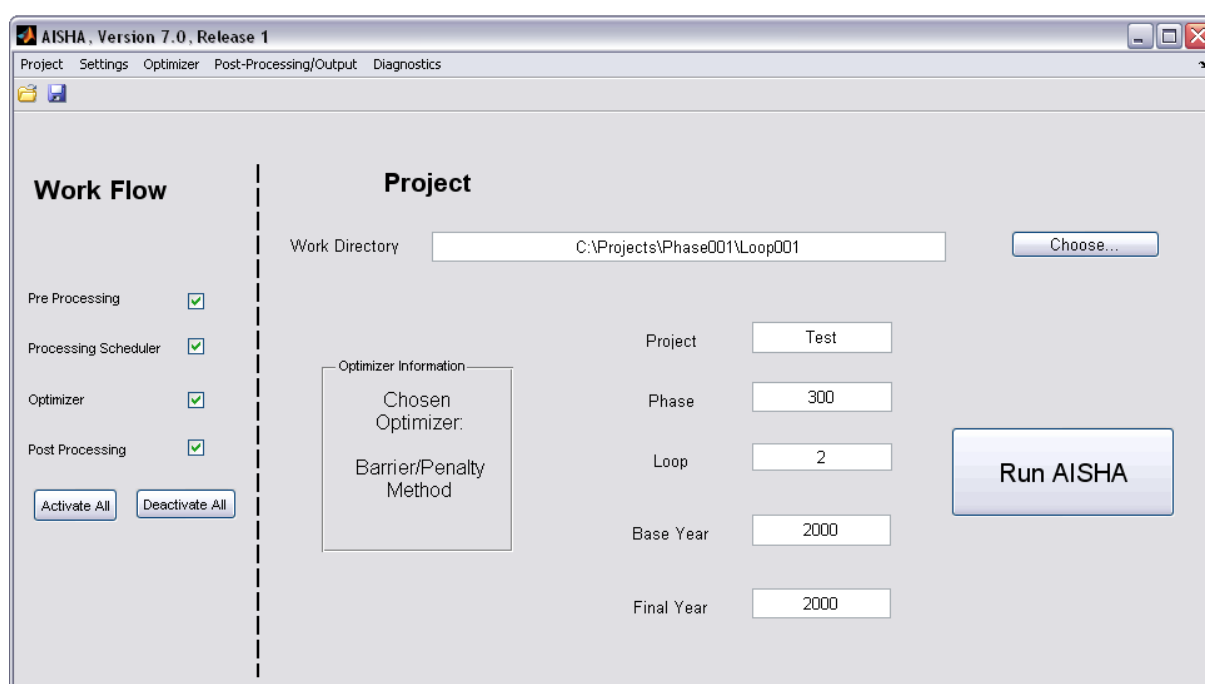

Figure 12: The AISHA-GUI

3.1.1   Programming Environment and Hardware

AISHA is programmed in Matlab and can be executed in Matlab version R2009b or later versions. Matlab offers a large variety of mathematical and particularly matrix operations and therefore offers an ideal environment for AISHA. AISHA, in its current version, can be executed in Matlab on Microsoft Windows.

3.1.2   Projects

The task of generating a series of contingency tables is called a *project* within AISHA. A project can be saved as a mat-file (the so-called project file) which contains all settings that the user made to define the project and the way it operates when generating the series of contingency

---

[19]The AISHA-acronym stands for *An Automated Integration System for Harmonised Accounts.*

tables.

Each project requires a *work directory* which is defined directly in the GUI. A work directory and its subdirectories contain all necessary input data for AISHA to complete the calculation process for a series of MRIO tables. Additionally, AISHA writes all generated output data files into the work directory and its subdirectories.

3.2   Input and Output Data

All of AISHA's data I/O is handled through the work directory. Some of the I/O files are Excel files. Supported Excel versions are 2003 or later. We will refer to later versions of Excel only and therefore only use the relevant file ending *xlsx*.

3.2.1   File Naming Convention

AISHA operates on a standard file naming convention. The AISHA-GUI uses this naming convention for its output files and most of its input files. The purpose of the naming convention is to provide a unique way of defining file names that offer detailed information about the contents of the file in arbitrarily large projects.

Before we can introduce the naming convention, the concepts of *phases* and *loops* have to be introduced. Throughout the development process of AISHA it became clear that certain projects might be run numerous times under different condidtions. These different condidtions could be for example different input data or a varying number of constraints. In order to provide some structure for these processes, AISHA supports phases and loops which are used to hierarchically structure the work directories of different projects. Loops are inferior to phases and it is recommended to sort work directories according to this structure. However, there is no strict definition for a phase or a loop.

Example 3.1. *(Phase/Loop Structure for Example 2.2)*
Consider there are numerous data sources for Example 2.2 and the first run of Example 2.2 is carried out only with half the available data. Then the we define the work directory for this first run as

<p align="center"><code>C:\Aisha\Phase001\Loop001</code></p>

and the for the second run which considers all the available data sources as

<p align="center"><code>C:\Aisha\Phase001\Loop002</code></p>

The AISHA-GUI supports this hierarchical structure. The phase and loop number are necessary user inputs of the AISHA-GUI (see Section 3.4.1) and are also part of the naming convention. Feasible values for the phase and loop are 1 to 999.

Definition 3.2. *(File Naming Convention for AISHA-GUI)*
The AISHA file naming convention defines a file name as a set of individual character strings divided by underscore characters according to the following convention:

<p align="center"><date str>_<project>_<countries>_<phase no>_<file id>_<loop no>_<arbitrary
string>.<file ext></p>

The elements of this file naming convention are

1. <date str>: The date string. This is an 8-digit number containing the date on which the file was created according to the following notation: *yyyymmdd*. Example: the 14th of May 2011 would be represented in the data string *20110514*.

2. <project>: The project name as it was defined in the by the user in the AISHA-GUI. See Section 3.4.1 for a detailed explanation.

3. <countries>: Countries affected by the information contained in the file. This is the three-digit UN country code that was already used in Example 2.2. Some files may only hold data that is related to one country, in which case the three digit UN country code would be used. For cases where all countries that are participating in the project are affected (for example the tree structure file, Section 3.5), this part is the string "AllCountries".

4. <phase no>: The phase number represented in three digits. For example "001" for phase 1. See Section 3.4.1 for for how to define the phase number of a project.

5. <file id>: The file identifier. This string defines what kind of data the file contains. This is the only part of the naming convention that is not defined by the project definition or the time at which the file is generated. Hence, this is the most important part of the naming convention.

6. <loop no>: The loop number represented in three digits. For example "002" for loop 2. See Section 3.4.1 for for how to define the loop number of a project.

7. <arbitrary string>: An arbitrary string the user may use for additional information. This string can stay empty.

8. <file ext>: The file extension. For example *txt* for text files or *xlsx* for Excel files.

Since all parts of the naming convention but the file identifier and the file extension are either defined by the project settings (Section 3.4.1), the amount of countries are affected by its contents, or the time at which the file was generated, we say that a file is named according to the naming convention using a certain file identifier. This uniquely defines the filename.

Example 3.3. *(File Naming Convention)*
Assume we have project called "TestProject" which is was defined as phase 1, loop 2. Then an Excel file named according to the naming convention with the file identifier "Results" generated on the 14th May 2011 would have the file name

$$20110514\_TestProject\_AllCountries\_001\_Results\_001.xlsx \ .$$

Clearly, a file having this file identifier would be expected to contain the results of a certain run.

3.2.2   AISHA-specific File Formats

The 8-DIM, 4-DIM and 1-DIM representations defined in Sections 2.1.1 and 2.1.3 can be used to define file formats that proved very useful for the use with AISHA.

Let $\mathbf{T}$ be a contingency table with I elements and let τ be the vectorisation of the contingency table. Let $\xi_{8,i}$ be the 8-tuple of the 8-DIM representation and $\tau_i = t_{\xi_{8,i}}$ the transaction value

of the contingency table $\mathbf{T}$ that is addressed by $\xi_{8,i}$. Then we can augment the 8-tuple by the value τ_i, yielding a 9-tuple

$$(x_{1,i}, x_{2,i}, x_{3,i}, x_{4,i}, x_{5,i}, x_{6,i}, x_{7,i}, x_{8,i}, \tau_i) \qquad \text{for} \qquad i = 1, \ldots, I \ . \tag{24}$$

This represents each element τ_i in an 8-dimensional sparse format (similar to the two dimensional sparse format used in Matlab). The same concept works for the 4-DIM and 1-DIM representation.

A total sparse representation of a contingency table $\mathbf{T}$ can be formulated by simply interpreting Equation 24 as a single line in a matrix which holds the sparse representations for all elements of $\mathbf{T}$. We obtain

$$\mathbf{T}_{sp8} := \begin{pmatrix} x_{1,1} & x_{2,1} & x_{3,1} & x_{4,1} & x_{5,1} & x_{6,1} & x_{7,1} & x_{8,1} & \tau_1 \\ x_{1,2} & x_{2,2} & x_{3,2} & x_{4,2} & x_{5,2} & x_{6,2} & x_{7,2} & x_{8,2} & \tau_2 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ x_{1,I} & x_{2,I} & x_{3,I} & x_{4,I} & x_{5,I} & x_{6,I} & x_{7,I} & x_{8,I} & \tau_I \end{pmatrix} \tag{25}$$

as an 8-DIM sparse representation of the whole contingency table $\mathbf{T}$. The matrix in Equation 25 has the dimensions $I \times 9$. The same concept can be applied to the 4-DIM and 1-DIM representation yield matrices with dimensions $I \times 5$ and $I \times 2$. The matrix given in Equation 25 can be stored in a tab-delimited text-file. This format of this file is called an *S8-file*. Using the same approach for the 4-DIM and 1-DIM representation, we obtain the *S4-file-* and the *S1-file*-formats.

The "S" stands for "sparse" and the number represents the amount of dimensions used in the sparse representation. Especially the S1-file format offers an economical way of storing an entire contingency table since only one dimension is required to uniquely identify the position of each transaction element. The disadvantage of these file formats is that the dimensions of the contingency table have to be available before this format can be used. The S1-, S4-, S8-file formats are used for numerous operations within AISHA.

In order to shorten file access time during I/O, these three formats are also supported in binary format. The structure is identical, but instead of storing the matrix from Equation 25 as a text-file, it is stored as a binary file. The file formats are then called B1-, B4-, and B8-file formats.

3.2.3   GUI-Inputs and Settings

The input data consists of three types of files:

1. A *tree structure file*. This file contains the structure of the series of contingency tables (amount of years, amount of regions, ...). This file is an Excel file which must have the file name **TreeStructure.xlsx**. It must be located in the work directory. The tree structure file is explained in detail in Section 3.5 (refer to Section 2.1.1 for an explanation of the structure of a contingency table).

2. A *constraints file*. This file holds all the information that is necessary to formulate all input data for the optimisation problem (Equation 23). This means, the constraints file holds all information to construct an initial estimate of the series of contingency tables, formulate the constraints matrix, define standard deviations for the given values, and to

apply the boundary conditions to the problem. The constriants file must have the file name **Constraints-$<yyyy>$.xlsx**, where $<yyyy>$ denotes the year that the constraints file is valid for. For example *Constraints-2000.xlsx* denotes the constraints file for the year 2000. It must be located in the work directory. The constraints file uses a meta-script language that is based on the 8-DIM representation from Definition 2.1 to address elements within the contingency tables. This meta-script is explained in detail in Sections 3.6 and 3.7.

3. *Data files.* The data files contain all necessary data to formulate the optimisation problem of Equation 23 (initial estimate τ_0, constraints data c, bounds l and u, and the standard deviation data σ_c and σ_τ. Supported data formats are comma-separated value files (csv-files) or the AISHA-specific file formats introduced in section 3.2.2. An arbitrary number of files can be used to supply all necessary data to the project. The commands within the constraints file define how the contents of the data files should be interpreted. Therefore, there are no restrictions whether the data files should contain single values, vectors or matrices of data. Data files can be stored in arbitrary locations on the system. The path to each individual file is defined within the constraints file.

3.2.4   Output Data

The resulting contingency table is written out into the work directory at the end of each time step by the AISHA-GUI. The user can choose in which format the data is written out (Section 3.4.5). There are two options for the output file format: Excel file or binary file format. Both these formats will be described in detail in Section 3.4.5.

3.3   Workflow

AISHA follows the methodology to construct a series of contingency tables as it was presented in Section 2. AISHA can be operated once all input files are available. AISHA's workflow is divided into four different steps which will be explained in detail within this section.

1. Pre-Processing

2. Processing Scheduling

3. Optimisation

4. Post-Processing

For a series on contingency tables, the workflow described in this section is carried out subsequently for each time step. In this case, all input files must be available for the whole series *before* commencing the calculations. The results of the completed year are automatically used as the intial estimate data for the subsequent years.

3.3.1   Pre-Processing

The pre-processing includes all operations that are necessary to prepare the contents of the data files so that they can be used in the processing scheduler. The pre-processing covers

- Ensure that the tree file and constraints file for the current year are available

- Read the constraints file contents and convert them to an internal Matlab-specific data structure that can be processed in the processing scheduler

- If previous years of a time series have already been completed: Locate the result files from the previous year and define these files as the initial estimate data files for the current time step.

3.3.2 Processing Scheduling

The processing scheduler completes all the work that is necessary to formulate the optimisation problem (Equation 23). Firstly, the precessing scheduler reads in the tree structure file and allocates memory space for a contingency table according to the tree structure data. Once the dimensions of the table are available and the memory is allocated, the processing scheduler reads and interprets the meta-commands that were loaded from the constraints file during the pre-processing. The meta commands hold all the information that is necessary to populate the contingency table with the initial estimate and to formulate the optimisation problem (Equation 23). In detail, the processing scheduler completes the following tasks by processsing the corresponding meta-commands:

1. Populate the contingency table by reading in the initial estimate data and placing the data in locations defined by the corresponding meta-command. For Example 2.2, this could mean that the Australian supply table is read from a csv-file and placed in the appropriate location in on the basic price sheet. The initial estimate is completed before any other commands are processed.

2. Read in the standard deviation values for the intial estimate of the contingency table.

3. After the initial estimate is completed, the processing scheduler vectorises the contingency table (and the standard deviations) according to Definition 2.4, yielding a vectorised table τ_0 and vectorised standard deviations σ_τ. Both vectors are input variables for the optimisation problem.

4. Construct each constraint according to Equation 19 and summarise the constraints in a constraints matrix $\mathbf{G}$ and a right-hand side vector c (see Equation 20). Matrix $\mathbf{G}$ and vector c are input variables for the optimisation problem.

5. Read in the standard deviation values of the right-hand side vector c and summarise them in vector σ_c. The vector σ_c is an input variable for the optimisation problem.

6. Define the lower and upper bounds for the elements in the contingency table and save them as vectors l and u. These vectors complete the set of input variables for the optimisation problem.

After the processing scheduler has finished the aforementioned tasks, all input variables are written out into a subdirectory of the work directory. The data is written out into binary format files which serve as input files for the optimisation algorithm. The fact that the data is written out and not passed on internally is based on the fact that external, non Matlab-based optimisation algorithms will be used in the future. Addtionally, the processing scheduler writes out a Matlab datafile (so-called mat-file) containing data to complete the post-processing.

3.3.3   Optimizer

AISHA offers two built-in optimisers to solve the optimisation problem (Equation 23).

1. A barrier-penalty optimiser based on Huang et al. (2008), and

2. An extension of the RAS-method that was developed by Lenzen et al. (2010a). This so-called KRAS optimiser is capable of balancing a contingency table under conflicting information.

3.3.4   Post-Processing

During the post-processing, the optimisation results are loaded from the result file of the optimiser and written out into result files. The user can choose between binary format (useful if the data is needed for further processing outside AISHA) or text format. The detailed order of operations during the post processing are

1. Load optimisation results from file. Load standard deviations of the result from file. Load mat-file to obtain necessary data for post-processing (mat-file was generated by the processing scheduler).

2. Invert the vectorisation of the contingency table. The necessary data for this operation was provided by the mat-file that was loaded during the previous step.

3. Write out results. Each valuation sheet of the contingency table is exported into an individual file.

4. In case there are further time steps to calculate as part of a time series: Write out the results into a B1-file (as introduced in Section 3.2.2). This result file will be used for Step 3 of the pre-processing as the initial estimate for the subsequent year.

3.4   Using the AISHA-GUI

Following the preparation of the input files (Section 3.2), the AISHA-GUI is used to define the project settings and to start the AISHA-algorithm. The AISHA-algorithm includes all operations described in the workflow (Section 3.3). The AISHA-algorithm is started by clicking on the "Run AISHA" button in the GUI (Figure 12). Once the AISHA-algorithm was started, the whole time series of contingency tables as defined by the project specifications will be calculated. Hence, all preparations and settings have to be completed before starting the AISHA-algorithm. In case of a time series project, all settings that were made within the AISHA-GUI are valid for each year of the time series.

Some of the options within the menus of the AISHA-GUI are switches to activate or deactive the option. Throughout this section, these menu options will be identified as *switch options*. The state of the switch is displayed by a check mark next to the option. A checked option is activated, an unchecked option is not.

3.4.1   GUI-Inputs and Settings

This section describes those inputs and settings of the AISHA-GUI that are located in the GUI itself and not in one of the menus. These settings define important framework settings for the

project as well as the workflow of the project. The AISHA-GUI will be described from left to right.

- **Disc icon** and **Folder icon** in the top left corner of the AISHA-GUI. Clicking on the folder opens a dialog to choose a project file, the disc icon opens a dialog to save a project file. The functionalities of these two icons are identical to the menu options *Load Project* and *Save Project* in the Project menu (Section 3.4.2).

- **Workflow.** The four checkboxes on the left hand side of the AISHA-GUI define which parts of the workflow are carried out when the AISHA-algorithm is initiated. The option to turn off individual parts of the workflow serve the purpose of time efficiency. Especially for large projects, each part of the workflow can take up a significant amount of time and it might be desirable to short-circuit certain parts of the workflow in order to save computing time. (Figure 12) shows a setting in which all four parts of the workflow are activated.
  In the current AISHA-GUI, the options *Pre-Processing* and *Processing Scheduler* must be activated at all times. The option to disable these workflow steps individually is reserved for future releases of AISHA.
  The *Optimiser* can be deactivated if for example no constraints are defined and AISHA is only used to construct a large contingency table from different data sources. In this case, the post-processing is able to use the data provided by the processing scheduler to complete the AISHA-algorithm and export the results.
  The option to deactivate the *Post-Processing* is helpful if for example the purpose of the current AISHA-run is to test a set of meta-commands. In this case it is useful to deactivate the options *Optimiser* and *Post-Processing* as the meta-commands are processed by the processing scheduler. **Note:** the post-processing has to be activated for time series projects. The post-processing ensures that the results of the current year are made available to the subsequent year as the initial estimate.

- **Activate All** and **Deactivate All** Buttons. These buttons check or uncheck all of the workflow checkboxes mentioned above.

- **Work Directory.** This text field displays the location of the work directory. The work directory can be defined by either typing the path into the text field, or by clicking on the *Choose...* button which opens the standard Windows user dialog to choose a directory.

- The box **Optimiser Information.** This box displays which optimiser was chosen. The optimiser is chosen in the optimiser menu (Section 3.4.4).

- **Project.** This text field is used to define the project name by simply typing the desired project name into the text field. The project name is a string of arbitrary length. It is used within the file naming convention introduced in Section 3.1.2.

- **Phase.** Type the phase number of the project into this text field (see Section 3.2.1). The project number is used for the file naming convention.

- **Loop.** Type the loop number of the project into this text field. The loop number is used for the file naming convention.

- **Base Year.** This field displays the initial starting year of the time series. The user is required to provide the base year in this field.

- **Final Year.** This field displays the final year of the time series. The user is required to provide the final year in this field. **Note:** The final year can be a smaller number than the base year. Time series projects that are directed backwards in real time are possible. A constraints file should be provided for each year of the time series. For each year that no constraints file is availble for, the result will be identical to the initial estimate.

- **Run AISHA** button. Click this button to start the AISHA-algorithm.

3.4.2   The Project Menu

The *Project* menu (Figure 13) offers a few standard options for the AISHA-GUI, but has no direct impact on how the AISHA-GUI operates. The individual options of this menu have the
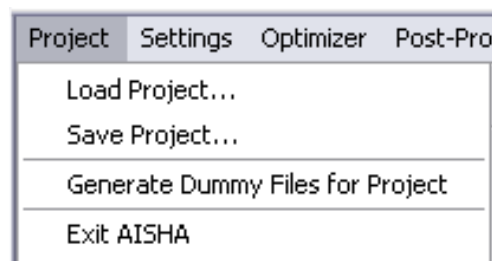

Figure 13: The Project menu of the AISHA-GUI.

following functionalities.

- **Load Project**. The option opens a dialog to choose a project file that was previously saved using the *Save Project* option. By loading a project file the AISHA-GUI is restored to the same state that it was in when the project was saved. This includes all settings and user inputs.

- **Save Project**. Saves the current state of the AISHA-GUI in a Matlab data file (mat-file). All user inputs, check box settings and menu settings are saved in the project file. Upon clicking this option, a standard Windows dialog to save files opens up. The default location for the project file is the work directory, the default file name defined according to the naming convention using the file identifier "Project". However, the location and name of the project file can be changed by the user without any restrictions.

- **Generate Dummy Files for Project.** By clicking this option, a set of blank input files are copied into the work directory. A dummy tree structure file and one constraints file for each year of the time series is produced. This option is intended to assist users with the correct format of input files.

- **Exit AISHA.** This option closes the program.

3.4.3   The Settings Menu

The *Settings* menu (Figure 14) offers a few additional general settings for the functionality of AISHA that are not directly linked to individual workflow steps. The individual options of this menu have the following functionalities.
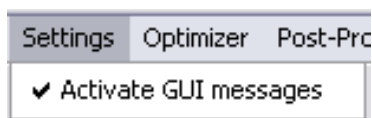
Figure 14: The Settings menu of the AISHA-GUI.

- **Activate GUI Messages** (switch option). The AISHA-GUI provides a number of standard outputs to the command window during its runtime. By activating this options, additional more detailed messages are displayed.

3.4.4   The Optimiser Menu

The *Optimiser* menu (Figure 15) defines the choice of optimiser for the project. All options
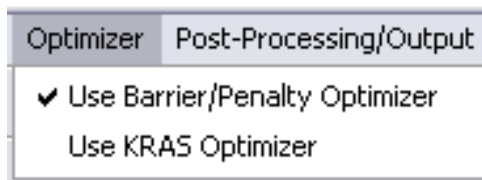
Figure 15: The Optimiser menu of the AISHA-GUI.

within this menu are switch options. Only one option can be activated at any time (the AISHA-GUI ensures that a maximum of one option is actiaved within this menu). The individual options of this menu have the following functionalities.

- **Use Barrier/Penalt Optimiser** (switch option). Activates a barrier-penalty optimiser for the current project.

- **Use KRAS Optimiser** (switch option). Activates the KRAS optimiser for the current project

The chosen optimiser is displayed in the text box *Optimiser Information* in the AISHA-GUI.

3.4.5   The Post-Processing Menu

The *Post-Processing* menu (Figure 16) defines which format is used for the output files of the project. All options are switch options and can be selected in any arbitrary combination. The

Figure 16: The Post-Processing menu of the AISHA-GUI.

individual options of this menu have the following functionalities.

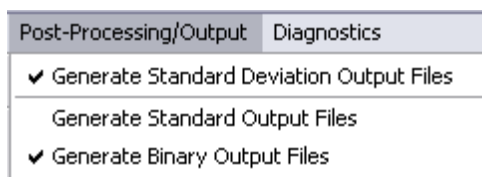- **Generate Standard Deviation Output Files** (switch option). When this option is activated, the AISHA-GUI writes out the standard deviations for the results of each valuation sheet each into a separate file in the work directory. The file is named according to the naming convention using the identifier "RsltSTDDEV-" followed by a four digit representation of the year, and the arbitrary string "Markup" followed by a 3-digit representation of the sheet number. For example

  20110510_StandardExample_AllCountries_300_RsltSTDDEV-2000_001_Markup002.bin

  for the second sheet.

- **Generate Standard Output Files** (switch option). Activating this options ensures that the AISHA-GUI saves the results for each year in one Excel file in the work directory. Each valuation sheet is saved in one separate sheet within the Excel file. The Excel file is named according to naming convention. The file identifier is "Results-" followed by a four digit representation of the year, and the arbitrary string "Markup" followed by a 3-digit representation of the sheet number.

- **Generate Binary Output Files** (switch option). Activating this options ensures that the AISHA-GUI saves the results and the standard deviation of the results for each year in binary files in the work directory. Each valuation is saved in a separate file. The files are named according to the naming convention using the same file identifiers as the first two options.

3.5   Tree Structure in AISHA

Within AISHA, the 8-DIM representation introduced in Definition 2.1 is used to address the contingency table elements. Each of the 8 dimensions have to be provided to AISHA in order to completely define the size of the contingency table within the program. A full definition of MRIO dimensions is called a *tree structure* within AISHA.

The terminology of the tree structure used in AISHA (Table 3) refers to the terminology used for the individual dimensions of the 8-DIM representation, yet it is not identical. The concept and terms used in for the tree structure were already introduced in Table 1.

In order to simplify the process of defining a tree structure for the user, a few requirements have to be met when defining the tree structure.

- Every row parent must have the same amount of row children

- Every column parent must have the same amount of column children

Example 2.2 shows a case where these requirements are not met initially. Japan has only two entities in row children (commodities and value added) but Austraila has three (industries, commodities and value added). In order to solve this problem, Japan has to be defined as having three row children, and the first row child simply has no row grandchildren. This would refer to Japan's industries not having any sectors.

The tree structure is defined by providing the maximum values for each of the 8 dimensions. Since row grandchildren and column grandchildren can have different amounts of sectors, the

sector amounts for each row grandchild and column grandchild must be defined individually (Table 4).

This representation of the tree structure is partly redundant. Due to the requirement that each parent must have the same amount of children, the column grandchildren and row grandchildren blocks have as many rows as they have parents and as many columns as they have children. The amount of row parents and column parents may differ (the same holds for row children and column children). Therefore the grandchildren blocks may differ in size for a tree structure.

Since the data for parents and children can be obtained from the size of the corresponding grandchildren block, only *Stacks*, *Sheets* and the blocks for *row grandchildren* and *column grandchildren* are provided to AISHA to define the tree structure. The tree structure data is provided to AISHA by the tree structure file. This file is an Excel file containing four sheets which is named according to the naming convention using the file identifier *TreeStructure*. These four sheets are named *Years (stacks)*, *Margins (sheets)*, *Rowgrandchildren* and *Columngrandchildren*.

The sheet named "Years (stacks)" (Figure 17) contains only the amount years for the contingency table in cell "A1" of the sheet. This number is given in Table 4 by the value "Stacks".

Figure 17: The "Years (stacks)" sheet of the tree structure file. This figure shows the corresponding sheet for Example 2.2. This data is identical to the data for "Stacks" in Table 4.

Recall that this number represents the years that are calculated at the same time. In a time series spanning more than one time step, this number still equals one, since each individual year

8-DIM	Aisha Terminology	Alternative Terminology
1	Stacks	Time; Number of years the contingency table covers
2	Sheets	Valuation; Number of valuation sheets
3	Row Parents	Region of flow origin (country, state, region, ...)
4	Row Children	Entity of flow origin (industry, commodity, satellite, ...)
5	Row Grandchildren	Sector of flow origin
6	Column Parents	Region of flow destination (country, state, region, ...)
7	Column Children	Entity of flow destination (industry, commodity, satellite, ...)
8	Column Grandchildren	Sector of flow destination

Table 3: Tree structure terminology in AISHA and alternative terms.

Stacks:	1					
Sheets:	5					
Row Parents:	4			Column Parents:	4	
Row Children:	3			Column Children:	3	
Row Grandchildren				Column Grandchildren		
Aus	3	3	1	Aus	3	3	1
Deu	2	0	1	Aus	2	0	1
Jpn	0	4	1	Jpn	0	4	1
RoW	1	0	0	RoW	1	0	0

Table 4: Tree structure for Example 2.2.

has to be processed before the calcualtions for the subsequent year can commence.

The sheet named "Margins (sheets)" (Figure 18) only contains the number of valuations of the contingency table in cell "A1" of the sheet. This value is given in Table 4 by "Sheets".
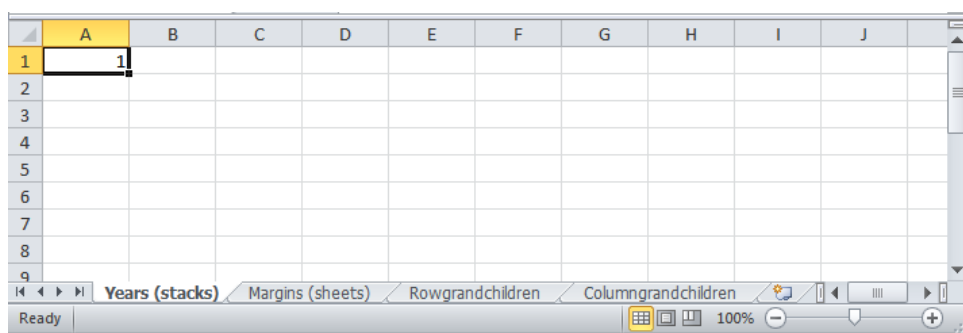

Figure 18: The "Margins (sheets)" sheet of the tree structure file. This figure shows the corresponding sheet for Example 2.2. This data is identical to the data for "Sheets" in Table 4.

The sheet "Rowgrandchildren" (Figure 19) contains the aforementioned row grandchildren block. The data has to be placed in the top left corner of the sheet. The row grandchildren block is given by "Row Grandchildren" in Table 4.

The sheet "Columngrandchildren" (Figure 20) contains the aforementioned column grandchildren block. The data has to be placed in the top left corner of the sheet. The column grandchildren block is given by "Row Grandchildren" in Table 4.

3.6   Addressing Contingency Table Entries Using Meta-Script

In this section we will present a simple command language that was developed to help users easily specify constraints on contingency tables. This *meta-command language* is designed to operate on the 8-DIM representation from Definition 2.1. The dimensions MRIO table that the meta-command language is operating on is defined by the tree structure (Section 3.5).

Figure 19: The "Rowgrandchildren" sheet of the tree structure file. This figure shows the corresponding sheet for Example 2.2. This data is identical to the data block under "Row Grandchildren" in Table 4.
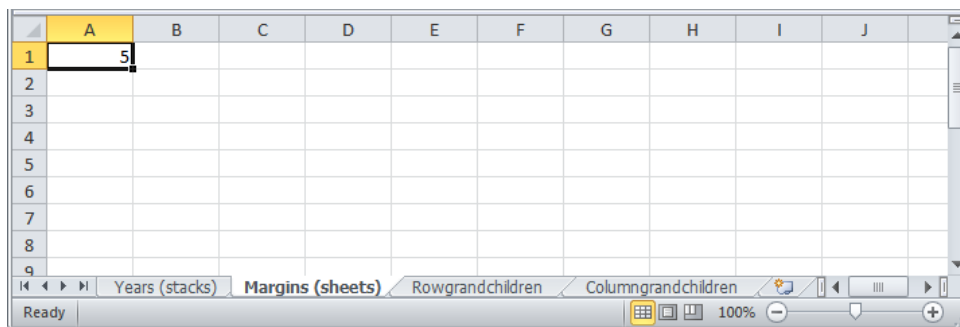

Figure 20: The "Columngrandchildren" sheet of the tree structure file. This figure shows the corresponding sheet for Example 2.2. This data is identical to the data block under "Column Grandchildren" in Table 4.

3.6.1 Basic Meta-Script Commands

For example, any entry in the MRIO introduced in Example 2.2 can be described by a vector in an 8-dimensional space. For example, a point $(1, 1, 2, 1, 3, 2, 1)$ addresses the entry for

- 2000 (1st year counted from base year 2000),

- Valuated in basic prices (valuation 1),

- Originating in Germany (country 2),

- As the industry (entity 1)

- 'Primary, manufacturing, and utilities' (industry 1),

- Destined for consumption in Japan (country 3),

- In the commodity (entity 2)

- 'Primary industries' (commodity 1).

This entry would be expressed in €.
A range of such entries can be conveniently expressed using basic meta-script commands as follows:

- $a : b$ denotes looping through all indices starting with a through to b;

- $a - b$ denotes summing over all indices starting with a through to b;

- $[a; b; c]$ denotes looping over (possibly disjunct) a, b, and c;

- $[a, b, c]$ denotes summing over (possibly disjunct) a, b, and c;

There exists a number of further meta-commands that address aggregated element ranges, aggregated data sources, and upper and lower bounds. The full meta-command suite will be published along with the release of the MRIO database in January 2012.

3.6.2 The Correlation Operator

The correlation operator $\sim$ works in the following way. Assume that index 4 is looped $1 : 3$ and index 6 is looped $1 : 3$, then this addresses $3 \times 3 = 9$ combinations of two indices. However, assume now that index 4 is looped $1 : 3$ and index 6 is looped $1 : 3 \sim 4$, then this addresses only 3 combinations of two indices, because index 6 is being tied to index 4, so that only coordinates $(1, 1)$, $(2, 2)$, and $(3, 3)$ are addressed. The same applies to examples such as $1 - 3$ (index 1) and $4, 9, 7 \sim 1$ (index 3), where entries are summed only over 3 coordinates $(1, 4)$, $(2, 9)$, and $(3, 7)$, and not over 9 coordinates as in $1 - 3$ (index 1) and $(4, 9, 7)$ (index 3).
A range of entries within the full 8-dimensional MRIO can be addressed for example as follows:

$$[1 : 1, 1 : 1, 3 : 3, 2 : 2, 1 - 4, 3 : 3, 2 : 2, 1 - 4] \tag{26}$$

could denote

- 2000 (1st year counted from base year 2000),

- Valuated in basic prices (valuation 1),

- Originating in Japan (country 3),

- As the commodity (entity 2)

    - 'Sum over all commodities' (commodities 1-4),

    - Destined for consumption in Japan (country 5),

    - In any of the commodities (entity 1)

    - 'Sum over all commodities' (commodites 1-4),

which sums up all entries within Japan's 4×4 commodity block in the basic price sheet. In contrast,

$$[1:1,1:1,3:3,2:2,1-4,3:3,2:2,4-4\sim 5] \tag{27}$$

could denote

    - 2000 (1st year counted from base year 2000),

    - Valuated in basic prices (valuation 1),

    - Originating in Japan (country 3),

    - As the commodity (entity 2)

    - 'Sum of all commodities' (commodities 1-4),

    - Destined for consumption in Japan (country 5),

    - In any of the commodities (entity 1)

    - 'Sum of all commodities, correlated to dimension 5' (commodites 1-4).

which only adds up the elements of the main diagonal of Japan's 4×4 commodity block in the basic price sheet. Depending on how many of the 8 operators are ':', '$-$', ';', ',', and '$\sim$', one meta-command can address a varying number of MRIO subsets.

3.6.3   Stacking up Constraints Blocks

Each constraint line in a meta-command list (see Section 3.6.1) results in a block component of the contraints matrix $\mathbf{G}$ that has the full number of columns, but varying numbers of rows, for example depending on whether a bock of elements are addressed one-by-one, or whether a sum is addressed, etc. The entire $\mathbf{G}$ can be assembled by simply looping through all constraint lines, and appending each block underneath what is already there.

3.6.4   Constructing c and σ_c

As with the constraint blocks (Section 3.8), the vectors c and σ_c are constructed by successively appending values and standard deviation of constraint block components.

    - LNx adds $\Sigma_{c,i} = xc_i$ (linear);

    - LG adds $\Sigma_{c,i} = \ln(c_i)$ (logarithm);

    - PWxEXy adds $\Sigma_{c,i} = xc_i^y$ (power / part of polynomial);

    - CNx adds $\Sigma_{c,i} = x$ (constant);

    - MXxMNy undertakes a regression of a data set in the form LGaCNb, on the basis of x being the relative standard error (RSE) of the maximum of the set, and y the RSE of the minimum of the data set (compare Lenzen et al. 2010b; Wiedmann et al. 2008).

3.7   Example: Meta-Script

The following example covers a range of typical constraints for Example 2.2.

3.7.1   The hierarchy trees

The hierarchy tree structure was defined as shown in Section 3.5.

3.7.2   The Meta-Command Sheet

All constraints are listed in a meta-command sheet (Figure 21).

In addition to the meta commands, this sheet contains switches ("Y" and "N") for turning on and off the constraints, a sequential numbering index, and an integer representing the number of parts in a compound constraint. Note that if a constraint contains more than one part, all parts must address the same number of elements, so that the constraint parts can be overlaid (see Section 3.6.3).

3.7.3   Defining the Initial Estimate

Before any constraints can be imposed on an MRIO table to be constructed, an initial estimate has to be defined. The meta-command below serves this purpose (Table 5), and we will now go through its constituents from left to right.

Columns 1-10:

1	Incl	#	Parts	Value	Pre-map	Post-map	S.E.	Pre-map	Post-map
IE	Y	1	1	I			E MX1;MN10;CN1e+3;		

Columns 11-19:

Coef 1	Years	Margins	from country	entity	of sector	to country	entity	of sector
A <data filename>	1	1	4	1	1:e	1	1	1:e

Table 5: A meta command defining a part of an initial estimate

1. Descriptive label, here "IE" for "Initial Estimate".

2. "Y" in order to turn this constraint on.

3. A sequential number, here constraint #1.

4. The number of parts in the entire command, here 1

5. A label "I" indicating that this command populates the initial estimate.

6. and 7. Empty; these fields will be explained later.

8. filename

9. and 10. empty; these fields will be explained later.

Figure 21: Example of a Meta-Command Input Sheet

11. "A" command signifying that what follows is a file reference to all values addressed in this meta-command. <data filename> is a placeholder for a path to a folder containing the data.

12. The initial estimate year "1"= 2000.

13. The valuation, here 1, or basic prices.

14. The country of origin, here 4, or the Rest of the World.

15. The entity of origin, here 1, or industries. Note that the tree structure (Section 3.5) does not allow a value larger than 1 in this case. Rest of the World is represented by a single sector. Any entry in this field larger than one will lead to this meta-command being ignored.

16. The sectors of origin, here 1:e, meaning "for all sectors from the first to the last". Once again, the tree structure specifies the RoW with only 1 sector, and AISHA interprets e=1, and sets this field to 1:1 = 1.

17. The country of destination, here 1, or Australia.

18. The entity of destination, here 1, or industries.

19. The sectors of destination, here 1:e=1:3, that is all Australian industry sectors (see Table 1).

The entire meta-command requires 3 data values, and indeed the data file <data filename> contains [1.34E+07 1.03E+07 1.25E+07].

The meta-command described in the previous paragraph dealt with trade from the Rest of the World to Australia, which represents one block in the schematic MRIO in Figure 3. Our example MRIO contains 139 such blocks:

- 1 domestic + 3 imports intermediate transaction blocks for 3 countries for 5 valuations (4x3x5=60)

- 1 domestic + 3 imports final transaction blocks for 3 countries for 5 valuations (4x3x5=60)

- 3 countries' export blocks to the Rest of the World (no domestic transactions) for 5 valuations (3x5=15)

- 3 countries' value-added blocks at basic prices (3)

- Australia's supply matrix at basic prices (1)

Hence our meta-command list contains commands that read in the data for the initial estimate in 139 steps from 139 separate files

Long meta-command lists are helpful because they are explicit and detailed in documenting what parts of the MRIO are being addressed by which data. However, for the experienced AISHA user such long lists may become a hinderance, and for such cases there exists the possibility to read in an initial estimate and its standard deviations using only one meta-command (Table 6). In this case, the initial estimate has been prepared in a previous run of AISHA and it is

Columns 1-7:

1	Incl	#	Parts	Value	Pre-map	Post-map
Initial Estimate (S4 structure)	Y	1	1	I		

Columns 8-11:

S.E.	Pre-map	Post-map	Coef 1
<filename>			S4 <IE data filename>

Table 6: Meta Command for S1 Initial Estimate. Columns 12-19 are empty for the S1 initial estimate meta command.

available as an S1- ,S4-, or S8-file (Section 3.2.2).

If instead of an S1, S4 or S8 command, a B1, B4 or B8 command is used, the sparse file is stored in binary format. An alternative to specifying standard deviations of the initial estimate through a formula (E command), these can be read in directly using the label 'S' (Table 7).

Once again, any sparse or binary sparse format can be used.

Columns 1-6:

1	Incl	#	Parts	Value	Pre-map
Standard Deviation of Initial Estimate (S1 structure)	Y	2	1	S	

Columns 7-11:

Post-map	S.E.	Pre-map	Post-map	Coef 1
	<filename>			S4 <SD data filename>

Table 7: Meta Command for S1 Standard Deviations of S1 Initial Estimate. Columns 12-19 are empty for the S1 standard deviation of initial estimate meta command.

4   Conclusions

This paper describes a tool called AISHA for constructing contigency tables. Contingency tables are essentially tables of counts. They are used to record variables (observations, subjects or measurements) into a number of categories. Examples for contingency tables that are often used to describe economies, and that are therefore also regularly published by national statistical agencies, are input-output tables (recording money transactions between industry sectors of economies), Social Accounting Matrices (recording money transactions between entities of a society) and satellite accounts (recording for example resource use, pollutant emissions and workforce population of industry sectors).

The entries of contingency tables are often subject to contraints. For example, most statistical bureaux use an optimisation method called RAS in order to balance their national input-output tables. Many researchers resort to optimisation tools in order to create their own specialised tables, often involving a substantial amount of money, time and human resources. To make these tasks easier, faster and less expensive is, in essence, the purpose of AISHA.

The first part (Section 2) focuses on the methodology of AISHA. Contingency tables are constructed from raw data which usually originate from a number of data sources. Additionally, elements of contingency tables may be subject to constraints, which usually lead to a reconciliation problem which can be interpreted as a undertermined optimisation problem. Constructing such a constrained table requires preparing the constraint information so that it can be interpreted by an optimisation algorithm, which then solves a optimisation problem. In order to achieve this goal, the original 8-dimensional representation of the contingency table is vectorised. Using the vectorised representation of the contingency table yields a matrix-by-vector notation for the constrained contingency table. Solving the optimisation problem then yields a contingency table which adheres to the contraints.

The second part (Section 3) presents the Matlab-based AISHA-tool which performs the methodology to construct a contingency table as presented in Section 2. The user has control over features that AISHA offers, AISHA's workflow and operational modes by using the graphical user interface (AISHA-GUI). As input data, the tool requires information about the dimensions of the contingency table (the so-called tree structure), data files that hold initial estimate data and constraint data, and a file containing commands to define the initial estimate and the contraints of the contingency table formulated in a special meta-language. These meta-commands address individual elements or blocks of elements within the contingency table by using the initial 8-dimensional structure. Using these input data, AISHA then constructs the initial estimate of the contingency table, formulates the constraints and prepares the input data for the optimiser. The tool then solves the optimisation problem using state-of-the-art optimisers and exports the results into a designated directory. Hence, AISHA faciliates all steps required to build a contingency table from initial estimate data and constraint optimisation. AISHA is the first tool of its kind for this kind of application. Examples of meta-commands applied to a highly-aggregated multi-region input-output (MRIO) table containing three countries are given in the appendix.

The version of AISHA presented in this paper is intended for stand-alone windows-based computer systems such as desktop computers or laptops. The tool is under constant development and current research includes designing a unix-based version of AISHA which has the capability to run in parallel on cluster-systems in order to improve efficiency, speed, and size of the resulting contingency table. The aim is to design a tool that is able to construct a time series of large contingency tables such as a series of global MRIO tables at a very high sector level according

to the methodology presented in Section 2. Due to the size and hardware requirements, the optimisation algorithm for large problems will be a parallel algorithm executed on super computer systems.

References

Allen, R. (1974). Some experiments with the RAS method of updating input-output coefficients. *Oxford Bulletin of Economics and Statistics*, 36:217–228.

Andrew, R. and Forgie, V. (2008). A three-perspective view of greenhouse gas emission responsibilities in New Zealand. *Ecological Economics*, 68:194–204.

Andrew, R., Lennox, J., and Peters, G. (2010). Greenhouse Gas Emissions Embodied in New Zealand's Trade: Methodology. In *Paper prepared for the 18th International Input-Output Conference*.

Andrew, R., Peters, G., and Lennox, J. (2009). Approximation and regional aggregation in multi-regional input-output analysis. *Economic Systems Research*, 21:311–335.

Bacharach, M. (1965). Estimating nonnegative matrixes from marginal data. *International Economic Review*, 6:294–310.

Bacharach, M. (1970). *Biproportional matrices & input-output change*. Cambridge University Press, Cambridge, UK.

Batten, D. and Martellato, D. (1985). Classical versus modern approaches to interregional input-output analysis. In *Australian Regional Developments No. 3-Input-output Workshop*, pages 39–66, Adelaide, Australia. Australian Government Publishing Service.

BBC News (2009). China seeks export carbon relief. `http://news.bbc.co.uk/2/hi/science/nature/7947438.stm`. Online article.

Bouwmeester, M. and Oosterhaven, J. (2008). Methodology for the construction of an international supply-use table. In *International Input-Output Meeting on Managing the Environment*, page in press, Sevilla, Spain.

Bullard, C. and Sebald, A. (1977). Effects of parametric uncertainty and technological change on input-output models. *Review of Economics and Statistics*, LIX:75–81.

Bullard, C. and Sebald, A. (1988). Monte Carlo sensitivity analysis of input-output models. *The Review of Economics and Statistics*, LXX:708–712.

Byron, R. (1978). The estimation of large social account matrices. *Journal of the Royal Statistical Society Series A*, 141(3):359–367.

Cole, S. (1992). A note on a lagrangian derivation of a general multi-proportional scaling algorithm. *Regional Science and Urban Economics*, 22:291–297.

Dalgaard, E. and Gysting, C. (2004). An algorithm for balancing commodity-flow systems. *Economic Systems Research*, 16(2):169–190.

Deming, W. and Stephan, F. (1940). On least-squares adjustments of a sampled frequency table when the expected marginal total are known. *Annals of Mathematical Statistics*, 11:427–444.

Dixon, R. (1996). Inter-industry transactions and input-output analysis. *Australian Economic Review*, 3/96:327–336.

Duchin, F. (1992). Industrial input-output analysis: implications for industrial ecology. *Proceedings of the National Academy of Science of the USA*, 89:851–855.

Duchin, F. and Steenge, A. (1999). Input-output analysis, technology and the environment. In van den Bergh, J., editor, *Handbook of Environmental and Resource Economics*, pages 1037–1059. Edward Elgar, Cheltenham, UK.

Evans, W. (1954). The effect of structural matrix errors on interindustry relations estimates. *Econometrica*, 22:461–480.

Forssell, O. (1998). Extending economy-wide models with environment-related parts. *Economic Systems Research*, 10:183–199.

Friedlander, D. (1961). A technique for estimating a contingency table, given the marginal totals and some supplementary data. *Journal of the Royal Statistical Society, Series A*, 124:412–420.

Gallego, B. and Lenzen, M. (2009). Estimating generalised regional input-output systems: A case study of Australia. In Ruth, M. and Davíðsdóttir, B., editors, *The dynamics of regions and networks in industrial ecosystems*, chapter 5, pages 55–77. Edward Elgar Publishing.

Gibbons, J., Wolsky, A., and Tolley, G. (1982). Approximate aggregation and error in input-output models. *Resources and Energy*, 22:203–230.

Gigantes, T. (1970). *The representation of technology in input-output systems*. North-Holland Publishing Company, Geneva, Switzerland.

Gilchrist, D. A. and St Louis, L. V. (1999). Completing input-output tables using partial information, with an application to canadian data. *Economic Systems Research*, 11(2):185–193.

Goicoechea, A. and Hansen, D. (1978). An input-output model with stochastic parameters for economic analysis. *AIIE Transactions*, 10:291–295.

Hanseman, D. and Gustafson, E. (1981). Stochastic input-output analysis: a comment. *Review of Economics and Statistics*, 63(A):468–470.

Hanseman, D. J. (1982). Stochastic input-output analysis: a simulation study. *Environment and Planning*, 14(A):1425–1435.

Heijungs, R. and Suh, S. (2002). *The computational structure of life cycle assessment*. Kluwer Academic Publishers, Dordrecht, Netherlands.

Heimann, M. and Reichstein, M. (2008). Terrestrial ecosystem carbon dynamics and climate feedbacks. *Nature*, 451:289–292.

Hertwich, E. and Peters, G. (2009). Carbon Footprint of Nations: A Global, Trade-Linked Analysis. *Environmental Science & Technology*, 43:6414–6420.

Hewings, G. and Madden, M. (1995). Social accounting: essays in honour of Sir Richard Stone. In Hewings, G. and Madden, M., editors, *Social and demographic accounting*, pages 1–14. Cambridge University Press, Cambridge, UK.

Hoekstra, R. (2010). (Towards) a complete database of peer-reviewed articles on environmentally extended input-output analysis. In *Paper prepared for the 18th International Input-Output Conference.*

Hondo, H., Sakai, S., and Tanno, S. (2002). Sensitivity analysis of total co_2 emission intensities estimated using an input-output table. *Applied Energy*, 72:689–704.

Huang, W., Kobayashi, S., and Tanji, H. (2008). Updating an Input-Output Matrix with Sign-preservation: Some Improved Objective Functions and their Solutions. *Economic Systems Research*, 20(1):111–123.

Kop Jansen, P. (1994). Analysis of multipliers in stochastic input-output systems. *Regional Science and Urban Economics*, 24:55–74.

Lahr, M. and de Mesnard, L. (2004). Biproportional techniques in input-output analysis: table updating and structural analysis. *Economic Systems Research*, 16(2):115–134.

Lave, L. B., Cobas-Flores, E., Hendrickson, C. T., and McMichael, F. C. (1995). Using input-output analysis to estimate economy-wide discharges. *Environmental Science & Technology*, 29:420A–426A.

Lecomber, J. (1975). A critique of methods of adjusting, updating and projecting matrices. In Allen, R. and Gossling, W., editors, *Estimating and Projecting Input-Output Coefficients*, pages 1–25. Input-Output Publishing Company, London, UK.

Lennox, J., Peters, G., and Andrew, R. (2010). Climate policy and emissions embodied in international trade. In Murray, J. and Wood, R., editors, *The Sustainability Practitioner's Guide to Input-Output Analysis*, pages 179–190. Common Ground, Champaign, USA.

Lenzen, M. (2000). Errors in conventional and input-output-based life-cycle inventories. *Journal of Industrial Ecology*, 4:127–148.

Lenzen, M., Gallego, B., and Wood, R. (2006). A flexible approach to matrix balancing under partial information. *Journal of Applied Input-Output Analysis*, 11& 12:1–24.

Lenzen, M., Gallego, B., and Wood, R. (2010a). Matrix balancing under conflicting information. *Economic Systems Research*, 21(1):23–44.

Lenzen, M., Wood, R., and Wiedmann, T. (2010b). 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. (1936). Quantitative input and output relations in the economic system of the United States. *Review of Economics and Statistics*, 18:105–125.

Leontief, W. (1941). *The Structure of the American Economy 1919-1939.* Oxford University Press, New York, NY, USA.

Leontief, W. (1974). Structure of the world economy. *American Economic Review*, LXIV:823–834.

Leontief, W. (1986a). *Input-Output Economics*, chapter An information system for policy decisions in a modern economy, pages 418–428. Oxford University Press, New York, NY, USA.

Leontief, W. (1986b). *Input-Output Economics*. Oxford University Press, New York, NY, USA.

Leontief, W., Chenery, H., Clark, P., Duesenberry, J., Ferguson, A., Grosse, A., Grosse, R., Holzman, M., Isard, W., and Kistin, H. (1953). *Studies in the Structure of the American Economy*. Oxford University Press, New York, NY, USA.

Leontief, W. and Ford, D. (1970). Environmental repercussions and the economic structure: an input-output approach. *Review of Economics and Statistics*, 52:262–271.

Leontief, W. and Ford, D. (1971). *Air pollution and the economic structure: empirical results of input-output computations*. North-Holland Publishing Company, Geneva, Switzerland.

Luthje, M., Pedersen, L., Reeh, N., and Greuell, W. (2006). Modelling the evolution of supraglacial lakes on the west greenland ice-sheet margin. *Journal of Glaciology*, 52:608–618.

Madsen, B. and Jensen-Butler, C. (1999). Make and use approaches to regional and interregional accounts and models. *Economic Systems Research*, 11(2):277–299.

Miller, R. and Blair, P. (1985). *Input-Output Analysis: Foundations and Extensions*. Prentice-Hall, Englewood Cliffs, NJ, USA.

Moran, D., Wackernagel, M., Kitzes, J., Heumann, B., Phan, D., and Goldfinger, S. (2009). Trading spaces: Calculating embodied ecological footprints in international trade using a product land use matrix (plum). *Ecological Economics*, 68:1938–1951.

Morrison, W. and Thuman, R. (1980). A lagrangian multiplier approach to the solution of a special constrained matrix problem. *Journal of Regional Science*, 20(3):279–292.

Munksgaard, J. and Pedersen, K. (2001). CO_2 accounts for open economies: producer or consumer responsibility? *Energy Policy*, 29:327–334.

Murray, A. T. (1998). Minimizing aggregation error in input-output models. *Environment and Planning*, A 30:1125–1128.

Nakićenović, N. and Swart, R. (2000). Special report on emissions scenarios. Technical report, Intergovernmental Panel on Climate Change, Geneva, Switzerland.

Nansai, K., Tohno, S., and Kasahara, M. (2001). Uncertainty of the embodied CO_2 emission intensity and reliability of life cycle inventory analysis by input-output approach. *Energy and Resources*, 22. (published by the Japan Society of Energy and Resources, in Japanese).

Oosterhaven, J., Stelder, D., and Inomata, S. (2008). Estimating international interindustry linkages: non-survey simulations of the asian-pacific economy. *Economic Systems Research*, 20:395–414.

Oosterhaven, J., Stelder, D., and Inomata, S. (2011). Constructing an environmentally-extended Multi-Region Input-Output table using the GTAP database. *Economic Systems Research*, 23:in press.

Paelinck, J. and Walbroeck, J. (1963). Etude empirique sur l'Èvolution de coefficients "inputoutput". *Economie Appliquée*, 16(1):81–111.

Park, S. (1973). On input-output multipliers with errors in input-output coefficients. *Journal of Economic Theory*, 6:399–403.

Park, S.-H., Mohtadi, M., and Kubursi, A. (1981). Errors in regional nonsurvey input-output models: analytical and simulation results. *Journal of Regional Science*, 21:321–339.

Peters, G. (2007). Efficient algorithms for Life Cycle Assessment, input-output analysis, and Monte-Carlo analysis. *International Journal of Life Cycle Assessment*, 12:373–380.

Peters, G. (2008). From production-based to consumption-based national emission inventories. *Ecological Economics*, 65:13–23.

Peters, G. and Hertwich, E. (2008). Co2 embodied in international trade with implications for global climate policy. *Environmental Science and Technology*, 42:1401–1407.

Polenske, K. R. (1980). *US multiregional input-output accounts and model.* D.C. Heath and Company, Lexington, MA.

Pyatt, G. (1991). Fundamentals of social accounting. *Economic Systems Research*, 3:315–341.

Quandt, R. (1958). Probabilistic errors in the Leontief system. *Naval Research Logistics Quarterly*, 5:155–170.

Rodrigues, J. and Domingos, T. (2008). *Carbon Responsibility and Embodied Emissions: Theory and Measurement.* Routledge, Florence, Kentucky, USA.

Roland-Holst, D. W. (1990). Interindustry analysis with social accounting methods. *Economic Systems Research*, 2(2):125–145.

Rueda-Cantuche, J., Beutel, J., Neuwahl, F., Mongelli, I., and Loeschel, A. (2009). A symmetric input-output table for eu27: Latest progress. *Economic Systems Research*, 21:59–79.

Sakai, S., Tanno, S., and Hondo, H. (2000). Uncertainty analysis for I/O analysis using perturbation method. In *The Fourth International Conference on EcoBalance*, pages 189–192, Tsukuba, Japan.

Schuster, U. and Watson, A. (2006). A variable and decreasing sink for atmospheric co_2 in the north atlantic. *Journal of Geophysical Research*, 112.

Serrano, M. and Dietzenbacher, E. (2010). Responsibility and trade emission balances: An evaluation of approaches. *Ecological Economics*, 69:2224–2232.

Smith, R., Weale, M., and Satchell, S. (1998). Measurement error with accounting constraints: Point and interval estimation for latent data with an application to U.K. Gross Domestic Product. *Review of Economic Studies*, 65:109–134.

Sonis, M. and Hewings, G. (1995). Matrix sensitivity, error analysis and internal/external multiregional multipliers. *Hitotsubashi Journal of Economics*, 36:61–70.

Sonis, M. and Hewings, J. (1989). Error and sensitivity input-output analysis: a new approach. In Miller, R., Polenske, K., and Rose, A., editors, *Frontiers of Input-Output Analysis*, pages 232–244. Oxford University Press, New York, USA.

Stephan, F. (1942). An iterative method of adjusting sample frequency tables when expected marginal totals are known. *Annals of Mathematical Statistics*, 13:166–178.

Stevens, G. and Trainer, G. (1980). Error generation in regional input-output analysis and its implications for nonsurvey models. In Pleeter, S., editor, *Economic Impact Analysis: Methodology and Applications*, pages 68–84. Martinus Nijhoff Publishing, Boston, USA.

Stone, R. (1986). Social Accounting: the state of play. *Scandinavian Journal of Economics*, 88(3):453–472.

Suh, S. and Huppes, G. (2005). Methods for Life Cycle Inventory of a product. *Journal of Cleaner Production*, 13:687–697.

Suh, S. and Nakamura, S. (2007). Five years in the area of input-output and Hybrid LCA. *International Journal of Life Cycle Assessment*, 12:351–352.

Tarancon, M. and Del Rio, P. (2005). Projection of input-output tables by means of mathematical programming based on the hypothesis of stable structural evolution. *Economic Systems Research*, 17(1):1–23.

Ten Raa, T. and Rueda-Cantuche, J. (2007). Stochastic analysis of input-output multipliers on the basis of use and make tables. *Review of Income and Wealth*, 53:318–334.

Ten Raa, T. and Steel, M. (1994). Revised stochastic analysis of an input-output model. *Regional Science and Urban Economics*, 24:361–371.

Thissen, M. and Löfgren, H. (1998). A new approach to sam updating with an application to egypt. *Environment and Planning Series A*, 30(11):1991–2003.

Tukker, A., Poliakov, E., Heijungs, R., Hawkins, T., Neuwahl, F., Rueda-Cantuche, J., Giljum, S., Moll, S., Oosterhaven, J., and Bouwmeester, M. (2009). Towards a global multi-regional environmentally extended input-output database. *Ecological Economics*, 68:1928–1937.

van der Ploeg, F. (1982). Reliability and the adjustment of sequences of large economic accounting matrices. *Journal of the Royal Statistical Society Series A*, 145(2):169–194.

van Vuuren, D. and Riahi, K. (2008). Do recent emission trends imply higher emissions forever? *Climatic Change*, 91:237–248.

Walter, K., Zimov, S., Chanton, J., Verbyla, D., and Chapin, F. (2006). Methane bubbling from siberian thaw lakes as a positive feedback to climate warming. *Nature*, 443:71–75.

Weber, C. (2008). Uncertainties in constructing environmental multiregional input-output models. In *International Input-Output Meeting on Managing the Environment*, Seville, Spain.

Weber, C. and Matthews, H. (2007). Embodied environmental emissions in U.S. international trade, 1997-2004. *Environmental Science and Technology*, 41:4875–4881.

West, G. (1986). A stochastic analysis of an input-output model. *Econometrica*, 54:363–374.

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., Lenzen, M., Turner, K., and Barrett, J. (2007). Examining the global environmental impact of regional consumption activities - part 2: Review of input-output models for the assessment of environmental impacts embodied in trade. *Ecological Economics*, 61:15–26.

Wiedmann, T., Lenzen, M., and Wood, R. (2008). Uncertainty analysis of the UK-MRIO model - Results from a Monte-Carlo analysis of the UK Multi-Region Input-Output model. Report to the UK Department for Environment, Food and Rural Affairs, London, UK. Stockholm Environment Institute at the University of York and Centre for Integrated Sustainability Analysis at the University of Sydney.

Wiedmann, T., Wilting, H., Lutter, S., Palm, V., Giljum, S., Wadeskog, A., and Nijdam, D. (2009). *Development of a methodology for the assessment of global environmental impacts of traded goods and services.* Environment Agency publication SCHO1009BRAM-E-P, Environment Agency UK and SKEP Network.

Wiedmann, T., Wilting, H. C., Lenzen, M., Lutter, S., and Palm, V. (2011). Quo Vadis MRIO? Methodological, Data and Institutional Requirements for Multi-Region Input-Output Analysis. *Ecological Economics*, submitted.

Wiedmann, T., Wood, R., Lenzen, M., Minx, J., Guan, D., and Barrett, J. (2010). The carbon footprint of the UK - Results from a Multi-Region Input-Output model. *Economic Systems Research*, 22:19–42.

Williams, E., Weber, C., and Hawkins, T. (2009). Hybrid approach to characterising and managing uncertainty in Life Cycle Inventories. *Journal of Industrial Ecology*, 13:928–944.

Yoshida, Y., H. Ishitani, R. M., Kudoh, Y., Okuma, H., Morita, K., Koike, A., Kobayashi, O., and Yoshioka, K. (2001). Evaluation of uncertainty in LCA based on input-output analysis. *Technology*, 8.

Zhou, X. (2009). How does trade adjustment influence national inventory of open economies? accounting embodied carbon based on multi-region input-output model. Technical report, Institute for Global Environmental Strategies (IGES), International Forum on Sustainable Asia and the Pacific, Shonan, Japan.

Zhou, X. (2010). Environmentally Extended Multi-Region Input-Output Model. In Murray, J. and Wood, R., editors, *The Sustainability Practitioner's Guide to Input-Output Analysis*, pages 167–178. Common Ground, Champaign, USA.

Zhou, X., Liu, X., and Kojima, S. (2009). Carbon emissions embodied in international trade: An assessment from the asian perspective. Technical report, Institute for Global Environmental Strategies (IGES), Hayama, Japan.

A Examples for Meta Commands

This section presents a number of meta commands and explains their impact on the MRIO table of Example 2.2. We need to use abbreviations for the sector names used in Example 2.2 which are defined in Table 8.

Country	Sector	Abbreviation
	Industry Sector 1: Primary, manufacturing and utilities	Aus I1
	Industry Sector 2: Trade and transport	Aus I2
	Industry Sector 3: other services	Aus I3
AUS	Commodity Sector 1: Primary, manufacturing and utilities	Aus C1
	Commodity Sector 2: Trade and transport	Aus C2
	Commodity Sector 3: other services	Aus C3
	Value Added Sector 1	Aus VA
	Industry Sector 1: Primary, manufacturing and utilities	Ger I1
DEU	Industry Sector 2: Services	Ger I2
	Value Added Sector 1	Ger VA
	Commodity Sector 1: Primary industries	Jpn C1
	Commodity Sector 2: Manufacturing and utilities	Jpn C2
JPN	Commodity Sector 3: Trade and transport	Jpn C3
	Commodity Sector 4: Other services	Jpn C4
	Value Added Sector 1	Jpn VA

Table 8: Abbreviations for sector names of Example 2.2

In this section we will showcase the use of AISHA on Example 2.2 by subsequently adding constraints to the MRIO. In each of the following sections, we will add constraints while keeping the constraints of all previous sections in the problem. The resulting MRIO of each run is presented by three sheets.

1. *Basic Price Table.* The basic price table of the MRIO.

2. *Margin Sheet.* The sum of the trade margin sheet and the transport margin sheet.

3. *Net Tax.* The sum of the taxes on products sheet and the subsidies on products sheet.

For each section, we will showcase one newly-introduced meta-command and mark the elements that are affected by the meta-command in the result tables.

All values in the tables of this section are displayed in billion US$.

A.1   Defining the Initial Estimate

The initial estimate is read from csv files. One of the meta-commands defining the initial estimate file is presented in Table 9.

Columns 1-10:

1	Incl	#	Parts	Value	Pre-map	Post-map	S.E.	Pre-map	Post-map
IE	Y	1	1	I			E LR-0.42125;CR10;CN1e+3;		

Columns 11-19:

Coef 1	Years	Margins	from country	entity	of sector	to country	entity	of sector
A <data filename>	1	1	2	1	1:e	1	1	1:e

Table 9: A meta command defining a part of the initial estimate

This meta-command addresses a certain block in the MRIO table (the elements of this block are shaded grey in Table 10) and defines the initial estimate for these elements to be read from the file defined in the "Coef 1"-column of the meta-command. The prefix "A" indicates that the data intended for the addressed elements has to be read from the file specified by <data filename>.

Basic Price Table (Sheet 1)

	Aus Ind			Com			**Deu** Ind		**Jpn** Com				**RoW**	**Aus** FD	**Deu** FD	**Jpn** FD
Aus I1				210.0	3.866	7.407										
Aus I2				10.91	167.6	14.08										
Aus I3				5.901	6.123	297.6										
Aus C1	73.93	21.43	24.61				0.461	0.114	3.917	9.883	0.574	0.253	39.78	76.68	0.271	1.004
Aus C2	25.72	36.53	28.43				0.008	0.029	0.002	0.081	0.284	0.151	7.343	94.32	0.003	0.428
Aus C3	22.44	34.66	64.54				0.034	0.074	0.006	0.175	0.570	0.305	14.88	235.7	0.014	0.864
Ger I1	0.306	0.091	0.141				426.7	218.2	0.493	5.452	0.493	1.119	388.1	0.290	404.8	5.435
Ger I2	0.304	0.443	0.548				231.0	529.3	0.007	1.325	0.712	0.487	194.3	1.076	1497	2.819
Jpn C1	0.063	0.009	0.017				0.166	0.057	21.21	54.54	36.55	21.11	2.799	0.033	0.066	76.54
Jpn C2	0.436	0.109	0.189				5.409	3.098	27.21	969.5	158.4	392.8	321.9	0.517	6.958	751.3
Jpn C3	0.233	0.349	0.431				0.165	0.557	22.93	231.9	236.2	241.1	12.12	0.837	0.068	1065
Jpn C4	0.455	0.682	0.841				0.778	1.721	15.63	302.5	277.2	536.5	24.28	1.631	0.178	2622
RoW	12.79	10.41	12.93				189.7	114.1	27.71	168.4	39.00	31.69		27.40	153.1	177.4
Aus VA	101.8	101.9	206.7													
Deu VA							504.1	1418								
Jpn VA									100.7	976.5	1046	2480				

Margin Sheet (Sheet 2 + Sheet 3)

	Aus Ind			Com			**Deu** Ind		**Jpn** Com				**RoW**	**Aus** FD	**Deu** FD	**Jpn** FD
Aus I1																
Aus I2																
Aus I3																
Aus C1	11.02	5.937	6.183				0.089	0.039	0.130	0.782	0.099	0.087	9.703	33.89	0.102	0.320
Aus C2	-11.70	-6.388	-6.416				-0.096	-0.046	-0.131	-0.800	-0.141	-0.111	-10.84	-34.49	-0.105	-0.387
Aus C3	0.019	0.065	0.067				0.006	0.008	0.001	0.018	0.042	0.023	1.134	0.458	0.003	0.067
Ger I1	0.023	0.011	0.021				50.61	50.15	0.021	0.153	0.027	0.198	21.14	0.031	189.9	0.156
Ger I2	-0.023	-0.011	-0.021				-50.61	-50.15	-0.021	-0.153	-0.027	-0.198	-21.14	-0.031	-189.9	-0.156
Jpn C1	0.000						0.000	0.000	6.145	19.74	16.46	8.409	0.001	0.000	0.000	54.00
Jpn C2	0.061	0.006	0.018				1.380	0.983	7.764	153.8	46.41	99.90	111.8	0.106	1.967	492.2
Jpn C3	-0.098	-0.060	-0.085				-1.468	-1.156	-14.03	-177.7	-65.81	-118.2	-113.7	-0.237	-1.977	-568.7
Jpn C4	0.036	0.054	0.067				0.089	0.173	0.052	1.827	0.996	2.615	1.898	0.130	0.010	5.890
RoW	-0.370	-0.301	-0.374				-5.490	-3.302	-0.802	-4.872	-1.129	-0.917		-0.793	-4.430	-5.132
Aus VA																
Deu VA																
Jpn VA																

Net Tax (Sheet 4 + Sheet 5)

	Aus Ind			Com			**Deu** Ind		**Jpn** Com				**RoW**	**Aus** FD	**Deu** FD	**Jpn** FD
Aus I1																
Aus I2																
Aus I3																
Aus C1	0.218	1.815	0.390				0.000	0.000	0.004	0.010	0.004	0.001		0.240	0.001	0.013
Aus C2	-0.180	-0.029	0.007				0.025	0.080	0.041	0.473	1.252	0.880		0.055	0.021	3.709
Aus C3	0.360	0.605	2.203				0.055	0.176	0.091	1.041	2.760	1.938		0.201	0.045	8.172
Ger I1	0.033	0.027	0.036				0.045	0.348	0.003	0.037	0.075	0.054		0.330	1.059	0.277
Ger I2	0.065	0.059	0.098				0.063	0.381	0.009	0.107	0.272	0.191		0.398	1.310	0.840
Jpn C1	0.003	0.003	0.004				0.002	0.005	-0.020	0.012	0.084	0.015		0.032	0.002	0.233
Jpn C2	0.061	0.057	0.062				0.025	0.073	0.027	0.082	0.585	0.162		0.771	0.092	2.542
Jpn C3	-0.047	-0.044	-0.073				-0.033	-0.106	-0.026	-0.031	-0.002	0.023		-0.108	-0.027	0.688
Jpn C4	0.097	0.091	0.151				0.071	0.222	0.045	0.118	0.218	0.725		0.238	0.062	2.181
RoW	0.287	0.165	0.168				0.094	0.130	0.007	0.065	0.038	0.041		-0.460	0.221	0.085
Aus VA																
Deu VA																
Jpn VA																

Table 10: The initial estimate for Example 2.2.

A.2 Using ":" and "-" meta-commands for balancing constraints

One of the meta-commands defining the balacing constraints that ensure the correct row/column-balance is presented in Table 11.

Columns 1-10:

1	Incl	#	Parts	Value	Pre-map	Post-map	S.E.	Pre-map	Post-map
m1 balancing Aus & Aus	Y	5	1	0			0		

Columns 11-19:

Coef 1	Years	Margins	from country	entity	of sector	to country	entity	of sector
1	1	2	1	1-e	1-e	1	[1;3]	1:e

Table 11: A meta command defining a balancing constraint

This meta-command presents the use of the ":" and "-" operators within balancing constraints. The aim of this constraint is to ensure that the column sums of domestic elements of country 1 (Australia) equal zero. The last column entries of the meta command (1, [1;3] 1:e) ensure that the sums defined in columns 14, 15, 16 are applied to each of the columns addressed by (1, [1;3] 1:e). Sine Australia has 3 industry sectors and 1 final demand sector, this meta-command defines four constraint lines g for the constraints matrix $\mathbf{G}$. The elements addressed by this meta-command are shaded grey in Table A.2.

Basic Price Table (Sheet 1)

	Aus Ind			Aus Com			Deu Ind		Jpn Com				RoW	Aus FD	Deu FD	Jpn FD
Aus I1				221.5	4.009	8.010										
Aus I2				11.63	175.7	15.39										
Aus I3				6.235	6.361	322.4										
Aus C1	70.09	20.10	23.29				0.464	0.115	3.639	9.164	0.545	0.242	37.74	72.75	0.257	0.952
Aus C2	24.80	34.85	27.37				0.008	0.029	0.002	0.076	0.274	0.147	7.086	91.02	0.003	0.413
Aus C3	20.75	31.71	59.57				0.034	0.072	0.005	0.158	0.528	0.285	13.77	218.1	0.013	0.800
Ger I1	0.288	0.084	0.133				426.7	217.0	0.455	5.020	0.465	1.063	365.7	0.274	381.4	5.121
Ger I2	0.288	0.415	0.518				232.3	529.3	0.006	1.227	0.675	0.466	184.1	1.020	1418	2.670
Jpn C1	0.064	0.010	0.017				0.180	0.062	21.21	54.44	37.36	21.75	2.859	0.034	0.067	78.17
Jpn C2	0.446	0.110	0.193				5.875	3.347	27.26	969.5	162.3	405.5	329.4	0.529	7.120	768.8
Jpn C3	0.233	0.345	0.429				0.175	0.588	22.43	226.4	236.2	243.0	12.11	0.837	0.068	1064
Jpn C4	0.451	0.668	0.831				0.819	1.800	15.17	293.0	275.0	536.5	24.07	1.617	0.177	2599
RoW	12.78	10.29	12.90				201.4	120.5	27.13	164.5	39.04	31.97		27.40	153.1	177.4
Aus VA	101.7	100.8	206.2													
Deu VA							535.1	1497								
Jpn VA									98.58	954.3	1047	2502				

Margin Sheet (Sheet 2 + Sheet 3)

	Aus Ind			Aus Com			Deu Ind		Jpn Com				RoW	Aus FD	Deu FD	Jpn FD
Aus I1																
Aus I2																
Aus I3																
Aus C1	11.34	6.124	6.264				0.089	0.039	0.130	0.782	0.099	0.087	9.703	33.96	0.102	0.320
Aus C2	-11.36	-6.192	-6.332				-0.096	-0.046	-0.131	-0.800	-0.141	-0.111	-10.84	-34.42	-0.105	-0.387
Aus C3	0.019	0.067	0.068				0.006	0.008	0.001	0.018	0.042	0.023	1.134	0.458	0.003	0.067
Ger I1	0.023	0.011	0.021				50.61	50.15	0.021	0.153	0.027	0.198	21.14	0.031	189.9	0.156
Ger I2	-0.023	-0.011	-0.021				-50.61	-50.15	-0.021	-0.153	-0.027	-0.198	-21.14	-0.031	-189.9	-0.156
Jpn C1	0.000	0.000	0.000				0.000	0.000	6.160	19.88	16.71	8.669	0.001	0.000	0.000	54.81
Jpn C2	0.061	0.006	0.018				1.380	0.983	7.784	154.8	47.11	103.1	111.8	0.106	1.967	499.5
Jpn C3	-0.098	-0.060	-0.085				-1.468	-1.156	-14.00	-176.5	-64.83	-114.5	-113.7	-0.237	-1.977	-560.3
Jpn C4	0.036	0.054	0.067				0.089	0.173	0.052	1.839	1.011	2.698	1.898	0.130	0.010	5.978
RoW	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Aus VA																
Deu VA																
Jpn VA																

Net Tax (Sheet 4 + Sheet 5)

	Aus Ind			Aus Com			Deu Ind		Jpn Com				RoW	Aus FD	Deu FD	Jpn FD
Aus I1																
Aus I2																
Aus I3																
Aus C1	0.466	2.136	0.417				0.000	0.000	0.005	0.013	0.004	0.001		0.240	0.001	0.013
Aus C2	0.000	0.000	0.041				0.053	0.168	0.081	0.918	2.493	1.765		0.055	0.021	3.709
Aus C3	0.361	0.601	2.227				0.059	0.188	0.090	1.026	2.785	1.972		0.201	0.045	8.172
Ger I1	0.059	0.051	0.076				0.089	0.422	0.007	0.079	0.188	0.135		0.330	1.059	0.277
Ger I2	0.178	0.163	0.273				0.098	0.423	0.025	0.287	0.770	0.545		0.398	1.310	0.840
Jpn C1	0.003	0.003	0.004				0.002	0.005	0.007	0.023	0.095	0.016		0.032	0.002	0.233
Jpn C2	0.102	0.094	0.125				0.058	0.173	0.053	0.107	0.697	0.176		0.771	0.092	2.542
Jpn C3	0.074	0.069	0.115				0.056	0.178	0.002	0.005	0.012	0.037		-0.108	-0.027	0.688
Jpn C4	0.099	0.092	0.154				0.077	0.240	0.045	0.118	0.221	0.743		0.238	0.062	2.181
RoW	0.287	0.163	0.167				0.100	0.137	0.007	0.063	0.038	0.041		-0.460	0.221	0.085
Aus VA																
Deu VA																
Jpn VA																

Table 12: The MRIO of Example 2.2 after introducing balancing constraints. The grey elements of the MRIO are affected by the balancing constraint in Table 11

A.3 Specifying the right-hand side

One of the meta-commands specifying the right hand side of a constraint is presented in Table 13.

Columns 1-10:

1	Incl	#	Parts	Value	Pre-map	Post-map	S.E.	Pre-map	Post-map
Define right-hand side	Y	88	1	<file name>	5.80E-04		E LR-0.22688;CR5;		

Columns 11-19:

Coef 1	Years	Margins	from country	entity	of sector	to country	entity	of sector
1	1	4-5	1	2	1-e	1	1	1-e

Table 13: A meta command defining a balancing constraint

This meta-command defines a sum over several elements of the MRIO (these elements are shaded grey in Table 14). The right-hand-side value for this sum is read from a file which is specified in the "Value"-column of the meta command.

Basic Price Table (Sheet 1)

	Aus Ind			Com			**Deu** Ind		**Jpn** Com				RoW	**Aus** FD	**Deu** FD	**Jpn** FD
Aus I1				221.5	4.009	8.010										
Aus I2				11.63	175.7	15.39										
Aus I3				6.235	6.361	322.4										
Aus C1	70.09	20.10	23.29				0.464	0.115	3.639	9.164	0.545	0.242	37.74	72.75	0.257	0.952
Aus C2	24.80	34.85	27.37				0.008	0.029	0.002	0.076	0.274	0.147	7.086	91.02	0.003	0.413
Aus C3	20.75	31.71	59.57				0.034	0.072	0.005	0.158	0.528	0.285	13.77	218.1	0.013	0.800
Ger I1	0.288	0.084	0.133				426.7	217.0	0.455	5.020	0.465	1.063	365.7	0.274	381.4	5.121
Ger I2	0.288	0.415	0.518				232.3	529.3	0.006	1.227	0.675	0.466	184.1	1.020	1418	2.670
Jpn C1	0.064	0.010	0.017				0.180	0.062	21.21	54.44	37.36	21.75	2.859	0.034	0.067	78.17
Jpn C2	0.446	0.110	0.193				5.875	3.347	27.26	969.5	162.3	405.5	329.4	0.529	7.120	768.8
Jpn C3	0.233	0.345	0.429				0.175	0.588	22.43	226.4	236.2	243.0	12.11	0.837	0.068	1064
Jpn C4	0.451	0.668	0.831				0.819	1.800	15.17	293.0	275.0	536.5	24.07	1.617	0.177	2599
RoW	12.78	10.29	12.90				201.4	120.5	27.13	164.5	39.04	31.97		27.40	153.1	177.4
Aus VA	101.7	100.8	206.2													
Deu VA							535.1	1497								
Jpn VA									98.58	954.3	1047	2502				

Margin Sheet (Sheet 2 + Sheet 3)

	Aus Ind			Com			**Deu** Ind		**Jpn** Com				RoW	**Aus** FD	**Deu** FD	**Jpn** FD
Aus I1																
Aus I2																
Aus I3																
Aus C1	11.34	6.124	6.264				0.089	0.039	0.130	0.782	0.099	0.087	9.703	33.96	0.102	0.320
Aus C2	-11.36	-6.192	-6.332				-0.096	-0.046	-0.131	-0.800	-0.141	-0.111	-10.84	-34.42	-0.105	-0.387
Aus C3	0.019	0.067	0.068				0.006	0.008	0.001	0.018	0.042	0.023	1.134	0.458	0.003	0.067
Ger I1	0.023	0.011	0.021				50.61	50.15	0.021	0.153	0.027	0.198	21.14	0.031	189.9	0.156
Ger I2	-0.023	-0.011	-0.021				-50.61	-50.15	-0.021	-0.153	-0.027	-0.198	-21.14	-0.031	-189.9	-0.156
Jpn C1	0.000	0.000	0.000				0.000	0.000	6.160	19.88	16.71	8.669	0.001	0.000	0.000	54.81
Jpn C2	0.061	0.006	0.018				1.380	0.983	7.784	154.8	47.11	103.1	111.8	0.106	1.967	499.5
Jpn C3	-0.098	-0.060	-0.085				-1.468	-1.156	-14.00	-176.5	-64.83	-114.5	-113.7	-0.237	-1.977	-560.3
Jpn C4	0.036	0.054	0.067				0.089	0.173	0.052	1.839	1.011	2.698	1.898	0.130	0.010	5.978
RoW	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Aus VA																
Deu VA																
Jpn VA																

Net Tax (Sheet 4 + Sheet 5)

	Aus Ind			Com			**Deu** Ind		**Jpn** Com				RoW	**Aus** FD	**Deu** FD	**Jpn** FD
Aus I1																
Aus I2																
Aus I3																
Aus C1	0.466	2.136	0.417				0.000	0.000	0.005	0.013	0.004	0.001		0.240	0.001	0.013
Aus C2	0.000	0.000	0.041				0.053	0.168	0.081	0.918	2.493	1.765		0.055	0.021	3.709
Aus C3	0.361	0.601	2.227				0.059	0.188	0.090	1.026	2.785	1.972		0.201	0.045	8.172
Ger I1	0.059	0.051	0.076				0.089	0.422	0.007	0.079	0.188	0.135		0.330	1.059	0.277
Ger I2	0.178	0.163	0.273				0.098	0.423	0.025	0.287	0.770	0.545		0.398	1.310	0.840
Jpn C1	0.003	0.003	0.004				0.002	0.005	0.007	0.023	0.095	0.016		0.032	0.002	0.233
Jpn C2	0.102	0.094	0.125				0.058	0.173	0.053	0.107	0.697	0.176		0.771	0.092	2.542
Jpn C3	0.074	0.069	0.115				0.056	0.178	0.002	0.005	0.012	0.037		-0.108	-0.027	0.688
Jpn C4	0.099	0.092	0.154				0.077	0.240	0.045	0.118	0.221	0.743		0.238	0.062	2.181
RoW	0.287	0.163	0.167				0.100	0.137	0.007	0.063	0.038	0.041		-0.460	0.221	0.085
Aus VA																
Deu VA																
Jpn VA																

Table 14: MRIO table of Example 2.2 with additional constraints. The grey shaded values are addressed by the meta-command given in Table 13