Routledge
Taylor & Francis Group

Economic Systems Research

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

A CYCLING METHOD FOR CONSTRUCTING INPUT–OUTPUT TABLE TIME SERIES FROM INCOMPLETE DATA

Manfred Lenzen , Maria Cecilia Pinto de Moura , Arne Geschke , Keiichiro Kanemoto & Daniel Dean Moran

Economic Systems Research, 2012, Vol. 24(4), December, pp. 413–432

A CYCLING METHOD FOR CONSTRUCTING INPUT–OUTPUT TABLE TIME SERIES FROM INCOMPLETE DATA

MANFRED LENZEN[a], MARIA CECILIA PINTO DE MOURA[b]*, ARNE GESCHKE[a], KEIICHIRO KANEMOTO[a,c] and DANIEL DEAN MORAN[a]

[a]*ISA, School of Physics, The University of Sydney, Australia;* [b]*Energy Planning Program, Graduate School of Engineering, Universidade Federal do Rio de Janeiro, Centro de Tecnologia, Bloco C, Sala 211, Cidade Universitária, Ilha do Fundão, Rio de Janeiro, Brazil;* [c]*Graduate School of Environmental Studies, Tohoku University, Sendai, Japan*

(*Received 1 September 2011; In final form 10 February 2012*)

There are a number of approaches for constructing time series of input–output tables. Some authors generate an initial estimate for a base year, and then serially estimate tables for subsequent years using the balanced prior-year table as an initial estimate. Others first generate a series of initial estimates for the entire period, and then balance tables in parallel. Current serial methods are affected by sudden leaps in the magnitude of table elements, which occur straight after a period of data unavailability. Current parallel methods require two complete tables for base and final years in the same classification, and therefore do not work under misaligned or incomplete data. We present a new method for constructing input–output table time series that overcomes these problems by averaging over alternate forward and backward sweeps across the time series period. We also solve the problem of hysteresis causing forecast and backcast table estimates to differ.

Keywords: Input–output tables, Time series, Matrix balancing, Forecasting, Backcasting

1. INTRODUCTION AND LITERATURE REVIEW

Perhaps the most widely publicised results so far derived from a time series of input–output tables is the examination of the UK's carbon footprint by Wiedmann et al. (2010) and Lenzen et al. (2010b), who demonstrated that – contrary to prior belief or myth – the UK's climate change responsibility had increased over the past decade, because emissions-intensive production was being outsourced to other countries, notably China. The political implications of communicating these findings to the public prompted the British Minister of the Environment to comment on whether the UK was in delusion over its emissions (BBC, 2008).

National statistical offices publish input–output tables over time. However, in virtually all cases, these tables do not adhere to a constant sector classification, and are not published every year. A notable exception are the tables issued by Statistics Denmark (1966–2007 at the time of writing; Statistics Denmark, 2011), which were used, for example, in a structural decomposition analyses (SDAs) by Wier (1998) and Wier and Hasler (1999).

---

*Corresponding author. E-mail: mcpmoura@gmail.com

ISSN 0953-5314 print; ISSN 1469-5758 online
http://dx.doi.org/10.1080/09535314.2012.724013

At the time of writing, the most detailed input–output time series constructed from incomplete and misaligned data existed for Australia, spanning the period 1974–2005. It was developed by Wood (2009), and subsequently applied to structurally decompose Australia's greenhouse gas emissions, to trace the evolution of its economic interconnectedness (Wood and Lenzen, 2009), and to document its material history (Wood et al., 2009). Similar time series were developed for many decades in order to publish SDAs – amongst others – for Austria (Skolka, 1989), China (Lin and Polenske, 1995), Taiwan (Chen and Rose, 1990; Wang, 1996; Chang and Lin, 1998), the US (Casler and Rose, 1998; Rose, 1999), India (Mukhopadhyay and Chakraborty, 1999), Netherlands (De Haan, 2001), Japan (Han, 1995; Kagawa and Inamura, 2001), the European Union (Alcántara and Duarte, 2004), Chile (Muñoz and Hubacek, 2008), the UK (Baiocchi and Minx, 2010) and Norway (Yamakawa and Peters, 2011).

These examples clearly show the relevance and usefulness of time series of input–output tables for policy- and decision-making. Notwithstanding their relevance, such time series are scarce, partly because of the substantial amount of labour and time required for their compilation, partly because of nearly unsurmountable challenges with respect to harmonising the often wildly varying product and industry classifications as well as currencies. Further, most time series are characterised by temporal gaps, that is, multiple years are missing. If a continuous, harmonised time series is required, analysts have to construct their own database from temporally incomplete and sectorally misaligned data. For example, plans for constructing continuous time series of global multi-region input–output tables were announced at the 18th International Input–Output Conference in Sydney (Lenzen et al., 2010a; Los and Stehrer, 2010). This article deals with challenges involved in developing such continuous, harmonised time series.

There are a number of methods for constructing time series of input–output tables. Some of these methods start with an initial estimate for a base year, and then serially estimate the table for each following year using the balanced prior-year table as an initial estimate. Other methods first generate a series of initial estimates for the entire period, and then balance each year in parallel. Whilst these methods differ with regard to their strengths and weaknesses (Temurshoev et al., 2011), they all are affected by at least one of the following two problems. The first problem is related to data missing for one or more intermediate years. Current serial methods are affected by sudden leaps in the magnitude of table elements, which occur straight after a period of data unavailability. We refer to the problem posed by such leaps as *hysteresis*. Current parallel methods suffer from the restriction that they require two complete tables for the base and final year in the same classification, meaning that they cannot be applied to sets of data that are inhomogeneous and/or incomplete for the base and final year. We will elaborate on these issues in the main part of this article.

For the example of the Brazilian Supply-Use Tables (SUTs) from 1970 to 2008, we present a new method for constructing input–output table time series that overcomes the problems described above by starting with only one initial estimate, and then averaging over alternate forward and backward sweeps across the entire time series period. In particular, we discuss how our method solves the problem of hysteresis that occurs when input–output transactions are dependent on whether they are the result of forecasting or backcasting over time.

Note that our claim is not to have developed a method that generates input–output tables that adhere more closely to the 'true' tables, however defined, as explored in Temurshoev et al.'s (2011) comparative study. In fact, our approach will work for any of the serial methods compared by Temurshoev et al.[1] The key innovation of our method is that it works hysteresis-free under circumstances in which current methods will either be affected by hysteresis or will not work at all because of excessive raw data inhomogeneity. Therefore, first, we do not test the performance of our method by comparing its output with known 'true' tables, because this is not the focus of our work. Second, we can compare our method neither with existing serial methods (because these were not devised to deal with hysteresis, and hence should not be compared on this basis) nor with existing parallel methods since these methods do not work on data sets with misaligned base- and final-year initial estimates.

In this study, we demonstrate the features of our method for the example of Brazil's SUTs between 1970 and 2008. A time series of Brazilian SUTs from 1970 to 1996 has previously been developed by Wachsmann et al. (2009), and applied to an SDA of Brazil's energy use. The difference between our work and that of Wachsmann et al. (2009) is that we extend the time series from 1996 to 2008, and that we use the 2005 product and industry classification, because the data supporting this intermediate year are more detailed than those supporting the 1996 classifications used by Wachsmann et al. (2009), and more detailed than those supporting 2000–2008 published SUTs. Note that the raw data for constructing this time series are such that superior data do not exist in the same classification for the base and final years, and the aggregation structure (1970) and incompleteness (2008) of data does not even allow constructing an initial estimate in the 2005 classification without further information. Such a situation exactly reflects the circumstances under which the method we propose has clear advantages over existing methods.

In the following section, we explain our methodology and data sources. In particular, we illustrate the problem of hysteresis and how it can be overcome by using a clear and simple example. Then, we explain the features of our approach using the Brazilian SUT time series. Finally, we draw some conclusions for future work of compiling input–output time series.

2. METHODOLOGY

2.1. Constrained Optimisation for Input–Output Table Balancing

The compilation of any input–output table requires the use of an optimisation method for table balancing, for example for the reconciliation of row and column totals. The approaches most often used for this task are variants of the RAS method, and various other optimisation methods (Robinson et al., 2001; Jackson and Murray, 2004; Lahr and de Mesnard, 2004; Huang et al., 2008; Temurshoev et al., 2011). These methods differ mainly by the type of objective function that is minimised. Any of these alternative methods could be used in the cycling method that is the main idea in this study.

Here, we balance the SUTs (vectorised as $\mathbf{a}$) by specifying an initial estimate (vectorised as $\mathbf{a}_0$), and applying the quadratic programming approach by Van der Ploeg (1988). External constraint information $\mathbf{c}$ (often called 'superior data') are linear functions $\mathbf{c} = \mathbf{G}\,\mathbf{a} + \varepsilon$ of the vectorised SUT entries $\mathbf{a}$, as well as disturbances ε that describe the constraint violation. Whilst any of the common optimisation approaches for table balancing would

[1] For example, RAS variants such as GRAS and KRAS, and also normalised, improved or weighted squared differences.

suit to demonstrate the principle of our cycling method, as mentioned above, we chose the quadratic programming approach because the disturbances allow effective handling of disparate, unaligned, conflicting and unreliable information (Lenzen et al., 2009), and because signs and zeros are not necessarily preserved. The sign- and zero-preservation inherent in the variants of the RAS balancing method (and other methods, see Huang et al., 2008) is undesirable because it does not allow account items such as net taxes and changes in inventories to switch signs, and it forces all variables connected to zero-valued constraints to zero without compromise.

Van der Ploeg extends $\mathbf{a}$ with the disturbances ε, to a compound unknown $\mathbf{p}$, distributed as

$$\mathbf{p} = \begin{pmatrix} \mathbf{a} \\ \boldsymbol{\varepsilon} \end{pmatrix} \sim \mathbf{D}\left[\begin{pmatrix} \mathbf{a}_0 \\ 0 \end{pmatrix}, \begin{pmatrix} \boldsymbol{\Sigma}_\mathbf{a} \\ \boldsymbol{\Sigma}_\mathbf{c} \end{pmatrix}\right] = \mathbf{D}\left[\mathbf{p}_0, \boldsymbol{\Sigma}\right] \quad (1)$$

with mean $\mathbf{p}_0 = [\mathbf{a}_0|0]$, and variance $\Sigma = [\Sigma_\mathrm{a}|\Sigma_\mathrm{c}]$. Note that the above formulation also caters for fundamental input–output balances (such as that each sector's gross input has to equal its gross output), where we write constraints as a difference that is forced to be zero, with the corresponding element in Σ_c also being zero, hence asking for an exact fit. Extending $\mathbf{C} = [\mathbf{G}| - \mathbf{I}]$, where $\mathbf{I}$ is the unity matrix, and assuming that all covariance terms in Σ vanish, the generalised quadratic problem becomes

$$\text{Minimise}\, f = (\mathbf{p} - \mathbf{p}_0)'\hat{\Sigma}^{-1}(\mathbf{p} - \mathbf{p}_0), \text{ subject to } \mathbf{Cp} = \mathbf{c}. \quad (2)$$

Setting up the Langrangean as $\mathcal{L} = (\mathbf{p} - \mathbf{p}_0)'\hat{\boldsymbol{\Sigma}}^{-1}(\mathbf{p} - \mathbf{p}_0) + \boldsymbol{\lambda}(\mathbf{Cp} - \mathbf{c})$, solving the first-order condition leads to analytical solutions $\boldsymbol{\lambda} = (\mathbf{C}\,\hat{\boldsymbol{\Sigma}}\mathbf{C}')^{-1}(\mathbf{Cp}_0 - \mathbf{c})$ and $\mathbf{p} = \mathbf{p}_0 - \hat{\boldsymbol{\Sigma}}\mathbf{C}\boldsymbol{\lambda}$; however, these do not guarantee any non-negativity that is usually imposed on all input–output transactions except subsidies and changes in stocks. We therefore add inequality constraints $l_i \leq p_i \leq u_i$ forcing the solution to lie within lower and upper bounds $l_i, u_i \in [-\infty, +\infty]$. The mixing of equality and inequality conditions precludes analytical solution, and requires sophisticated numerical solvers.

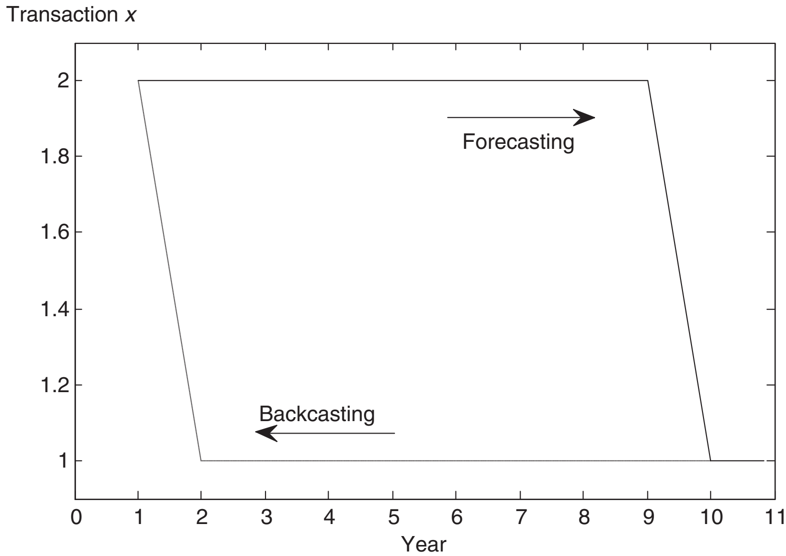
2.2. The Problem of Hysteresis in Forecasting and Backcasting

When estimating input–output time series, the superior data are usually affected by temporal gaps, that is, information on input–output transactions is usually not available for all years within the time series period. In order to overcome this problem, researchers have estimated transactions that are not supported by data in 1 year on the basis of information available for nearby years.

National statistical offices often use a prior-year table as an initial estimate for compiling an updated table. If such a procedure is carried out annually such as by Statistics Denmark (2011), a continuous time series is obtained. However, in almost all countries, input–output tables are issued with multiple-year gaps between them. If a continuous time series is desired, the question arises as to which value unknown intermediate-year transactions should assume.

Temurshoev et al. (2011) give a comprehensive comparison of eight methods that can be used to update series of input–output or SUTs. Six of these methods (RAS and squared differences variants) start with an initial estimate for a base year, and then serially estimate the table for each following year using the balanced prior-year table as an initial estimate. The remaining two methods (EUKLEMS and Euro) first generate a series of initial estimates

FIGURE 1. Hysteresis effect of successive fore- and backcasting of an input–output transaction.



Notes: x fixed to $x_1 = 2$ and $x_{10} = 1$ only for years 1 and 10. Data are assumed missing for year 11, and the year-11 solution follows the year-10 solution.

for the entire period, and then iteratively balance each year in parallel. We will first explain a problem that affects any of the *serial* methods.

Assume an input–output transaction x for which superior data exists in year 1 and in year 10, with $x_1 = 2$ and $x_{10} = 1$. Assume further that no data support the estimation of x in years $2, \ldots, 9$ and in year 11. If an initial estimate is set up for year 1, and years $2, \ldots, 11$ are enumerated by progressing the solution from year 1 forward in time, then $x_{2,\ldots,9} = 2$. The solution in year 10 is always $x_{10} = 1$, because it is fixed by a constraint representing the superior data point. The solution in year 11 is always $x_{11} = 1$, because similar to the forecasting progression across years $2, \ldots, 9$, there are no data available to alter the value for x. If in contrast an initial estimate were set up for year 10, and years $9, \ldots, 1$ were enumerated by progressing the solution from year 10 backward in time, then $x_{2,\ldots,9} = 1$.[2] In other words, the solution for the intermediate years depends on whether forecasting or backcasting is employed. Successive fore- and backcasting would result in a hysteresis curve (Figure 1), and hence we refer to this problem as *hysteresis*.

The choice between fore- and backcast solutions does not only arise out of having to deal with temporal gaps. For example, in our case, we needed to make a choice about which product and industry classifications to use for our SUT time series. We chose the 2005 classifications, but we could have chosen the ones for 1970. In the latter case, we would have taken 1970 as our only initial estimate year, and all subsequent-year information as superior data supporting constraints for a forward-progressing optimisation

[2] The case of missing data for year 0 would be identical to that of missing data in year 11, so we do not treat it explicitly.

procedure. But actually, we selected 2005 as our initial estimate year and, when constructing the tables for the period 1970–2005, we use all prior-year information as superior data supporting constraints for a backward-progressing optimisation procedure. Hence, if we only ever considered one optimisation sweep of the time series period, the magnitude of the resulting SUTs would have a significant and undesirable dependence on the choice of initial estimate and classification. Obviously, such arbitrariness needs to be avoided.

One way of avoiding the sudden leaps in the values of table elements is to pre-define a series of initial estimates, one for each year, that is characterised by smooth inter-year transitions of table values, and then to balance each of those initial estimates separately according to the constraints imposed by the data for the respective year. The EUKLEMS and Euro methods investigated by Temurshoev et al. (2011) are two examples for such a *parallel* approach. Similarly, Wood (2011) uses regression techniques in order to inter- and extrapolate incomplete initial estimate data over time.[3] Without going into details of these parallel methods, we stress at this point that all of the above parallel approaches require the initial estimate of the base and final years to exist in the same sector classification. Essentially, this requirement prevents the usage of such approaches for any application where the available data do not support the construction of such homogeneously classified initial estimates.[4] For example, in our case, some of the 1970 superior data are an aggregate of elements to be estimated in the 2005 classification, and the 2008 data are incomplete. None of the parallel methods could operate using these data sets without further information.

In order to circumvent the restrictions imposed by incomplete and misaligned initial estimate data whilst at the same time ensuring smooth (that is hysteresis-free) inter-year transitions, Tarancon and Del Rio (2005) apply the stable structural evolution hypothesis, and use adjacent-year values in order to formulate bounds as constraints for unknown input–output coefficients, which stay active during balancing. However, since these bounds are fixed, it may happen during matrix optimisation that they conflict with other constraints.

---

[3] Wood finds that, depending on which years' data are available, and the trends inherent in these data, some inter- and extrapolation functions yield unrealistic results. Wood settles on a sign-preserving exponential regression, but requires additional caps in order to limit regression results to realistic values. Whilst such a regression approach smoothes the initial estimate, it does not respect the constraints in the subsequent balancing procedure. Hence, smooth trends of input–output entries in the initial estimate may be overridden and destroyed by constraints posed by non-smooth data and by balance constraints.

[4] In principle, one could think of ways to interpolate initial estimates from inhomogeneous base- and final-year data; however, catering for all conceivable data gap structures would be far from trivial. Essentially, a linear interpolation between any two known neighbouring-year values would appear desirable and feasible. However, in practice, this is not possible because temporal gaps in data supporting real input–output time series are far from simple. First, there may be more than one supporting data point for a transaction in a particular year, and some of the supporting data may be aggregates relating to a sub-sum of the table to be estimated. In such complex situations of data availability, one necessarily has to choose which data to base the initial estimates on, and exclude any data that overlaps and/or conflicts with the chosen data (Wood, 2011). This means that initial estimates thus constructed are either labour-intensive (if based on a large number of inhomogeneous supporting data) or relatively arbitrary and conflicting (if based on a small number of inhomogeneous supporting data). For example, Baiocchi and Minx (2010) use weighted averages of neighbouring tables in order to interpolate missing intermediate input–output tables in a multi-region input–output time series constructed around the UK by Wiedmann et al. (2010). In order to substitute missing information at the beginning and end of the analysis period, they use constant technology coefficients of the earliest and latest available years. However, they qualify that 'these are strong assumptions, which can only be justified by the lack of global input-output data'.

This circumstance forces Tarancon and Del Rio to introduce a set of supplementary variables representing such incompatibilities, which in turn are subjected to an optimisation procedure, until the input–output table time series has a feasible solution. For large-scale and highly automated software platforms such as described by Lenzen et al. (2010a), and also used in this study, such a procedure is problematic, mostly because it would require a sensitivity analysis for every temporal gap.

One could re-formulate Tarancon and Del Rio's conditions by introducing inter-temporal constraints into the optimisation problem, such as $x_n - (x_{n-1} + x_{n+1})/2 \leq \text{tol}$, where tol is a tolerance value for the deviation of any year-n transaction value x_n from the average $(x_{n-1} + x_{n+1})/2$ of its adjacent-year values x_{n-1} and x_{n+1}. Such inter-temporal constraints would, however, turn N separate optimisation problems for N years into one single optimisation problem that is at least N times larger. Given present requirements for computer memory and run-time (Lenzen et al., 2010a), a system of inter-temporal constraints is clearly prohibitive for large-scale applications.

These examples may give the reader an impression of the challenges involved in achieving inter-temporal continuity from severely incomplete and inhomogeneous data. In the following section, we examine an approach that employs alternate fore- and backcasting of input–output tables across the time series period. Similar to Tarancon and del Río, we constrain table elements not only by imposing values of superior data for any current year, but at the same time, and as much as possible, align each current year's table elements with neighbouring-year elements. However, rather than using explicit constraint for the temporal alignment, we accomplish smooth inter-year transitions by averaging over alternate forward and backward sweeps across the entire time series period.

2.3. Forecast and Backcast Cycling

The basic idea behind fore- and backcast cycling is to take year-wise averages of the forecast and backcast optimisation run solutions. More specifically, instead of simply being prior-year solutions as described in Section 2.2, initial estimates $\mathbf{a}_0^{c,y}$ for year y in cycle c are weighted sums of all prior-year $y - z$ and all prior-cycle $c - d$ solutions $\mathbf{a}^{c-d,y-z}$

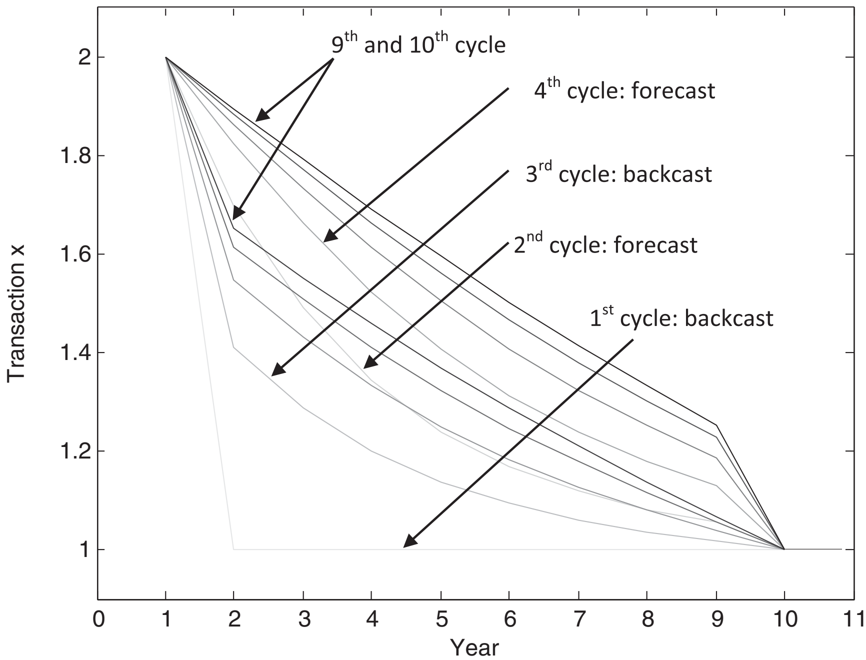
$$\mathbf{a}_0^{c,y} = \sum_{d=0}^{c-1} \sum_{z=0}^{y-y_0} \mathbf{w}(d, z)\mathbf{a}^{c-d,y-z} \tag{3}$$

with d and z being delay indices, and the weights $\boldsymbol{w}$ being normalised through $\sum_{d=0}^{c-1} \sum_{z=0}^{y-y_0} \mathbf{w}(d, z) = 1$. Note that 'prior-year' can mean 'earlier' or 'later' year, depending on whether c is a forecasting or a backcasting cycle.

In the following section, we illustrate the effects of cycling for the simple example in Figure 2, and for the special case of $\mathbf{w} = \begin{pmatrix} 0 & \alpha \\ 1 - \alpha & 0 \end{pmatrix}$, where α is a same-cycle, prior-year weight, and 1–α is a prior-cycle, same-year weight. In other words, every initial estimate is a weighted sum between the prior-year solution in the actual cycle, and the same-year solution in the previous cycle. Since two adjacent cycles always proceed in opposite directions, this choice of $\boldsymbol{w}$ effectively facilitates an ongoing averaging between forecast and backcast cycles, and thus leads to a continuous convergence to a unique time series.

Following a backcast as in figure (lightest grey in Figure 2), the first averaging occurs in year 2 of the following forecast (second cycle in Figure 2), where $x_{\text{prior-year}} = 2$,

FIGURE 2. Ten-cycle hysteresis for $\alpha = 0.7$.



Notes: The first cycle (lightest grey) is a backcast from year 10 to year 1, which then becomes averaged with a subsequent forecast, and so on. Successive cycles are distinguished by increasing shades of grey.

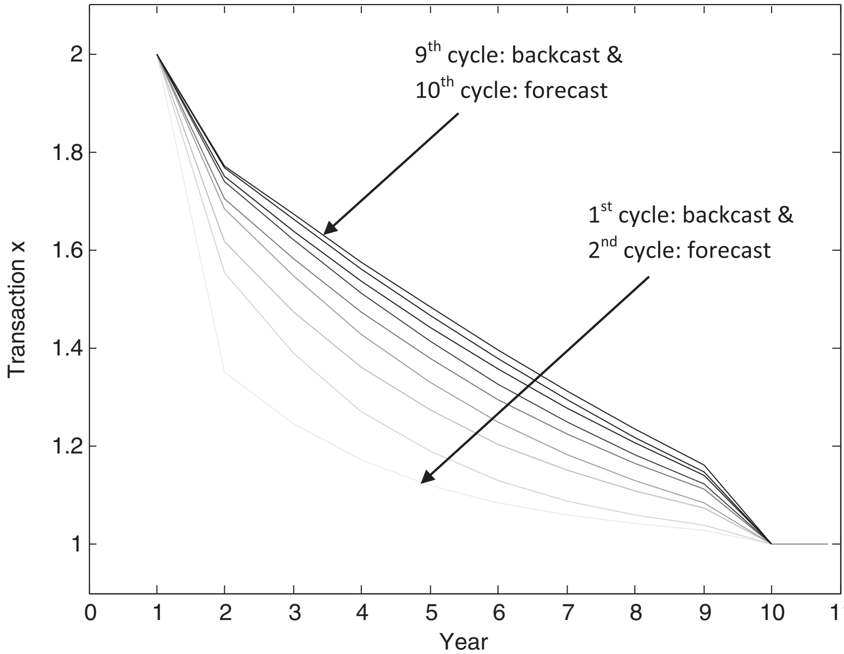
$x_{\text{prior-cycle}} = 1$, and the weighted average becomes $0.7 \times 2 + 0.3 \times 1 = 1.7$. In the next year, the averaging leads to $0.7 \times 1.7 + 0.3 \times 1 = 1.49$, and so on. The next backcast (third cycle in Figure 2) then decreases the average again, because the prior-year transactions start in year 10 with a value of $x_{\text{prior-year}} = 1$. Thus, the alternate fore- and backcasting continues, and the hysteresis curve becomes narrower and more symmetrical (see darker shades in Figure 2).

Seeing that forecasts and backcasts remain distinct even after 10 cycles (Figure 2), it is clear that neither of them represents a unique representation of the transaction time series. However, since the hysteresis curve becomes more and more symmetrical, two-cycle moving averages nicely converge towards a unique time series (Figure 3). In this study, we take the final (i.e. most stable) two-cycle moving average $\bar{\mathbf{a}}^{N,y} = 0.5(\mathbf{a}^{N-1,y} + \mathbf{a}^{N,y})$ as our input–output table solution.

Setting the same-cycle, prior-year weight to $\alpha = 0$ means that the initial estimate is always and only taken from the same-year solution of the prior cycle. This means that each subsequent cycle will reproduce the hysteresis profile of the first cycle. This can be seen in Figure 4 where the 9th/10th-year cycle profile at $\alpha = 0$ is identical to the profile of the initial backcast in Figure 1.

On the other extreme, setting the same-cycle, prior-year weight to $\alpha = 1$ means that the initial estimate ignores any prior cycle and only ever considers same-cycle, prior-year values. This means that subsequent cycles will go through the same hysteresis as in figure

FIGURE 3. Ten-cycle hysteresis for $\alpha = 0.7$ as in Figure 2, but showing moving averages $\bar{\mathbf{a}}^{N,y}$ over two adjacent cycles.



Notes: Successive adjacent-cycle averages are distinguished by increasing shades of grey. The adjacent-cycle average converges to a unique time series.

over and over again, and the moving two-cycle average never changes from the step shape shown in Figure 4 at $\alpha = 1$. Intermediate values will produce gradual transitions as in Figure 3. In this study on Brazil's SUTs, we use $\mathbf{w} = \begin{pmatrix} 0 & \alpha \\ 1-\alpha & 0 \end{pmatrix}$ with $\alpha = 1/2$.

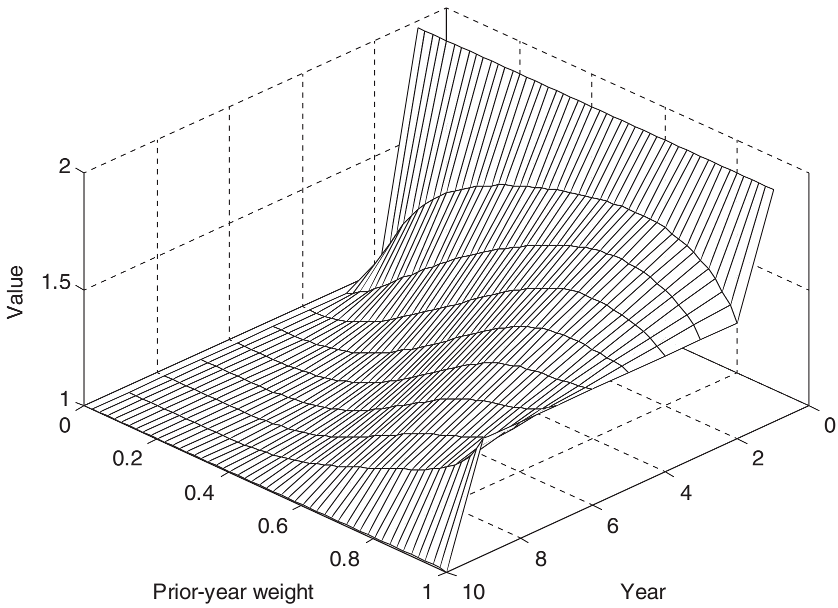
The cycling scheme we apply to Brazil's SUTs proceeds as in Table 1. We experimented with up to $N = 11$ cycles, but ultimately used fourth- and fifth-cycle iterates to construct the final SUT time series, because the differences between the SUTs generated during these cycles were already much smaller than 1% of the SUT means.

3. DATA SOURCES

3.1. Brazilian SUTs

Brazil's first SUT set was published in 1979 by the Brazilian Institute of Geography and Statistics (IBGE, 2010). It was compiled for the year 1970, on the basis of a nationwide economic and demographic census. The tables were updated every 5 years until 1990, following annual updates until 1996 (see details in Table 2). Following the recommendations made in the United Nations (UNs) Systems of National Accounts, the IBGE developed a new methodology that permitted synchronisation between the SUTs and the country's System of National Accounts. The 1990 and subsequent tables reflect this new approach, setting

FIGURE 4. Ninth/10th-cycle average $\bar{\mathbf{a}}^{10,y}$ as in Figures 2 and 3, but for varying same-cycle, prior-year weights α.



Notes: We have omitted year 11 from this graph because the year-11 solution is identical to the year-10 solution.

TABLE 1. Fore- and backcast cycling in the estimation of Brazil's SUT time series.

Cycle	Direction	Cycle-year label						Increment
1	Backcast		1C-1971	...	1C-2003	1C-2004	1C-2005	−1
2	Forecast	2C-1970	2C-1971	...	2C-2003	2C-2004		1
3	Backcast		3C-1971	...	3C-2003	3C-2004	3C-2005	−1
4	Forecast	4C-1970	4C-1971	...	4C-2003	4C-2004		1
5	Backcast		5C-1971	...	5C-2003	5C-2004	5C-2005	−1
6	Forecast	6C-1970	6C-1971	...	6C-2003	6C-2004		1
7	Backcast		7C-1971	...	7C-2003	7C-2004	7C-2005	−1
8	Forecast	8C-1970	8C-1971	...	8C-2003	8C-2004		1

Notes: 'C′ = cycle. A new cycle is started whenever the base year or the final year is reached.

them apart from the earlier tables, which were not integrated into the System of National Accounts (IBGE, 2008).The IBGE has published a complete SUT time series from 1990 to 2008 (IBGE, 2011); however, in this database, the use table is only available in purchasers' prices, and no margins and tax matrices are available to derive a basic price use table. Since our aim is to estimate a SUT time series expressed in basic prices, we directly utilised from this data set only the supply matrix, value added and gross output. We used the IBGE use and final demand matrices to place constraints on the relative proportions of the basic-price use and final demand matrices to be estimated. This procedure is equivalent to assuming that

TABLE 2. Sector details for Brazil's SUTs.

Year	Products	Industries	Final demand categories	Value-added categories
1970	160	90	11	12
1975	261	127	15	18
1980	136	91	5	6
1990–1996	80	43	5	8
1997–1999	80	42	7	14
2000, 2005	110	55	6	8
2001–2004, 2006–2008	110	56	7	11

taxes and margin are a constant proportion of basic prices. The most recent SUT completely available in all valuations (basic and purchasers' prices) at the time of writing was that for 2005. We therefore use this SUT for our initial estimate, while the SUTs corresponding to all previous years are used as superior data to constrain the solutions of our optimisation runs. For years with SUTs unavailable,[5] we fix total gross domestic product (GDP) to its current-year value using data in IPEA (2010).

Each SUT database contains multiple tables, starting with the supply and use matrices, value added, final demand, import matrices, sectoral participation and other coefficient matrices. Except for the 1990–2008 SU tables, a reconciliation matrix provides a detailed breakdown of total supply and use valued in purchasers' prices, into basic prices and all taxes and margins.

3.2. Harmonising Industry and Product Classifications Using Concordance Tables

The changes which took place over the years in the methodology used to implement the Brazilian SUTs are reflected in the changing product and industry classification listed in Table 2. In order to be able to use these data to formulate constraints on the SUT series to be estimated, the data classifications and SUT series classifications have to be related to each other using concordance matrices. The translation of these classifications leads to the first serious challenge in constructing a harmonised SUT time series. We stress at this stage that these translations do not lead to base- and final-year initial estimates in the desired 2005 classification (enabling one of the parallel approaches described by Temurshoev et al., 2011), because first, part of the 1970s data are too aggregated to achieve this, and second, the 2008 data are incomplete. Our method works even with such data that are grossly insufficient for constructing base- and final-year initial estimates. The concordances described in this section are set up for the purpose of formulating constraints that relate whatever imperfect data are available, to the table elements to be estimated.

The industry and product classifications used in the 1970, 1975 and 1980 SUTs are mostly but not always more detailed than those used in 1985, 1996, 2000 and 2005. Nevertheless, and considering also that the more recent SUT database is incomplete for our purposes, we chose to cast our SUT time series in terms of the most recent (2000 and 2005) product and industry classification, hoping that further updates would adhere to this standard. Compared

[5] 1971–1974, 1976–1979, 1981–1984 and 1986–1989.

to the entire time period, the two recent classifications are of intermediate detail (Table 2). As a consequence, the 1970, 1975 and 1980 tables are mostly aggregated before being imposed as constraints on the harmonised SUT. In contrast, the 1990s' tables have to be used as constraints on sub-sums of the harmonised SUT (compare Lenzen et al., 2006). In both cases, one needs a means of translating or, mapping, input–output transactions from one classification into another. This is conveniently achieved using concordance matrices.

Each concordance matrix maps the classification categories from the previous years (1970, 1975, 1980 and 1996) into the more recent 2005 classification categories, by placing a 1 in a cell where two classes overlap, and a 0 otherwise (compare Lenzen et al., 2010a, Section 2.2.3.2 and Appendices 1 and 2). Once again, the aggregation structure of the 1970s data does not allow constructing an initial estimate for 1970 in terms of the 2005 classification (as required by existing parallel methods) and as a result, some 1970s data points relate to more than one 2005-classified table element. In total, 16 concordance matrices were constructed, one for each group of common-classified years (Table 2), each in turn subdivided into four separate concordances for industries, products, value-added categories and final demand categories.

The Brazilian National Economic Activity Classification (CNAE; UN, 2010) was used as the underlying guideline for the necessary aggregation and disaggregation procedures. This classification is compatible with the third revision of the International Standard Industrial Classification, recommended by the UNs Statistical Commission for the purpose of harmonising global economic information. The recent classification for 2000 and 2005 adopted by the IBGE is based on the CNAE classification.

In order to determine the most appropriate and consistent mappings, certain assumptions had to be made. First, products and industries grouped within a certain CNAE hierarchy have been kept together wherever possible. For instance, according to the CNAE classification, 'Pig Iron' is a group (24.1) in the 'Metallurgy Division' (24), and appears in the 1970 matrix. However, in the 2000/2005 industry classification, there is no separate classification for pig-iron, only a 'Steel Manufacturing and Steel Products' class. A concordance between the 1970 and 2000 matrices was hence made by including the 1970 pig-iron in the 2000 'Steel Manufacturing and Steel Products' industry, rather than in the 'Other Metal Products' industry, because the latter is listed in CNAE as a different division.

Second, a harmonisation problem occurs regarding fictitious 'dummy' sectors that the IBGE includes in various years' SUTs in order to compensate for the fact that certain products, such as financial services, do not constitute intermediate consumption of the productive sectors. As a result, it is assumed that such products are consumed by a dummy sector, such as a fictitious financial sector, the total production of which adds up to zero (Carvalheiro, 1998). The problem is that such dummy sectors do not appear in the 2000/2005 classification. To complicate matters further, it is not always clear whether as sector is a dummy sector or not, for the term 'dummy' is not consistently part of the classification label, such as in the recycling sector in the 1975 industry classification. Whenever we located such sectors in earlier classifications, we attempted to reclassify these to match the most similar productive sector in the 2000/2005 classification.

3.3. Tackling Currency Changes

The second challenge is posed by Brazil's changing currencies, which are a reflection of the history of the Brazilian economy during the 35 years which span the chosen period of

TABLE 3. Comparison of Brazil's GDPs in actual currency units for selected years as in the input–output tables, and in 2009 Reais according to IPEA (2010).

Year	Input–output table currency		GDP (actual currency)	GDP (2009R$)	Conversion factor
1970	10^6 Cr$	Cruzeiro	Cr$ 189,865	R$ 692,171,670	3645.60
1975	10^6 Cr$	Cruzeiro	Cr$ 843,886	R$ 1,118,327,860	1325.21
1980	10^6 Cr$	Cruzeiro	Cr$ 11,690,557	R$ 1,583,388,910	135.44
1985	10^6 Cr$	Cruzeiro	Cr$ 1,180,380	R$ 1,686,793,580	1429.03
1990	10^6 Cr$	Cruzado (Novo)	Cr$ 31,759,185	R$ 1,851,108,470	58.29
1991	10^6 Cr$	Cruzeiro	Cr$ 165,786,498	R$ 1,870,202,370	11.28
1992	10^6 Cr$	Cruzeiro	Cr$ 1,762,636,611	R$ 1,861,470,110	1.06
1993	10^6 CR$	Cruzeiro Real	CR$ 38,767,062	R$ 1,948,310,500	50.26
1994	10^3 R$	Real	R$ 325,617,200	R$ 2,052,240,400	6.30
1995	10^3 R$	Real	R$ 661,309,085	R$ 2,142,884,410	3.56
2000	10^3 R$	Real	R$ 1,179,482,000	R$ 2,367,127,260	2.01
2005	10^3 R$	Real	R$ 2,147,239,000	R$ 2,715,609,450	1.26

analysis. In the late 1980s and the early 1990, Brazil passed through periods of very high inflation rates (for example, 81% in March 1990). In order to combat such high inflation, the Brazilian currency was re-defined six times between 1970 and 1996, so that most of the input–output tables are expressed in different currencies, rendering the harmonisation of tables rather difficult. We converted all input–output tables into constant 2009 Reais by benchmarking the input–output tables against constant-price GDP data in IPEA (2010) (Table 3, compare with Wachsmann et al., 2009).

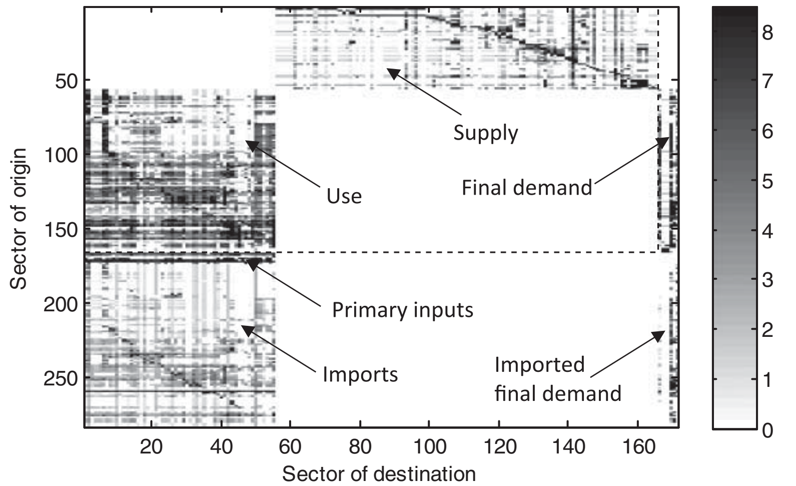
4. RESULTS

After completing N fore- and backcast cycles across the entire time series period, we have accumulated a data set $a_i^{n,y}$ of SUTs, vectorised across sectors i, and one each for every cycle n and every year y. From this data set, we compute the final (i.e. most stable, see Figure 3) vectorised two-cycle averages $\bar{\mathbf{a}}^{N,y}$ as our time series solution.

Figure 5 shows the classical SUT structure for table $\bar{\mathbf{a}}^{5,1996}$ of 1996. The shades of grey represent the logarithm of input–output transactions ranging from 1 2009R$ to 1 billion 2009R$. As expected, the use table, the final demand table, primary inputs and the supply table's diagonal hold most of the dominant transactions. Imports into intermediate and final demand are several orders of magnitude smaller. In the use and import matrices, the discernible off-diagonal represents sector-internal transactions, with the shape of the off-diagonal determined by the product-to-industry concordance. Note that even though supply, use and import matrices are rectangular (110 products $\times$ 55 industries), the entire intermediate SUT block (dashed outline in Figure 5) is square (165 $\times$ 165 sectors).

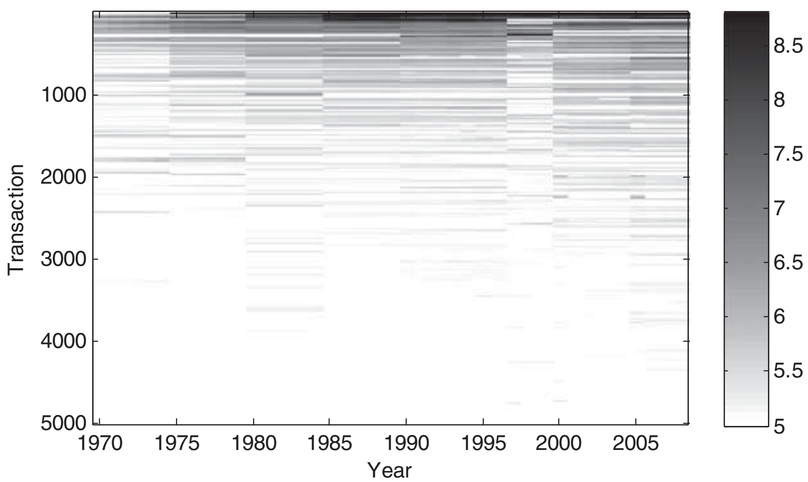
Visualising an input–output table time series is not as straightforward because of the added dimension of time. In this study, we take the final two-cycle average $\bar{\mathbf{a}}^{N,y}$ and compute for each element the sum across all years $\bar{\mathbf{A}}^N = \sum_y \bar{\mathbf{a}}^{N,y}$. We then sort $\bar{\mathbf{A}}^N$ in descending order, yielding a sequence of sectors $\bar{\mathbf{A}}^{N*}$ ranked in terms of their overall importance throughout the entire time series period. We then re-arrange the year-wise data $\bar{\mathbf{a}}^{N,y}$ so it is sorted

FIGURE 5. Topographic map of Brazil's 1996 SUT expressed in the 2005 classification.



Notes: Units are logarithms of transactions in 2009R$.

FIGURE 6. Topographic map of a time series of Brazil's vectorised ninth-cycle moving-average SUTs $\bar{\mathbf{a}}^{9,y*}$.

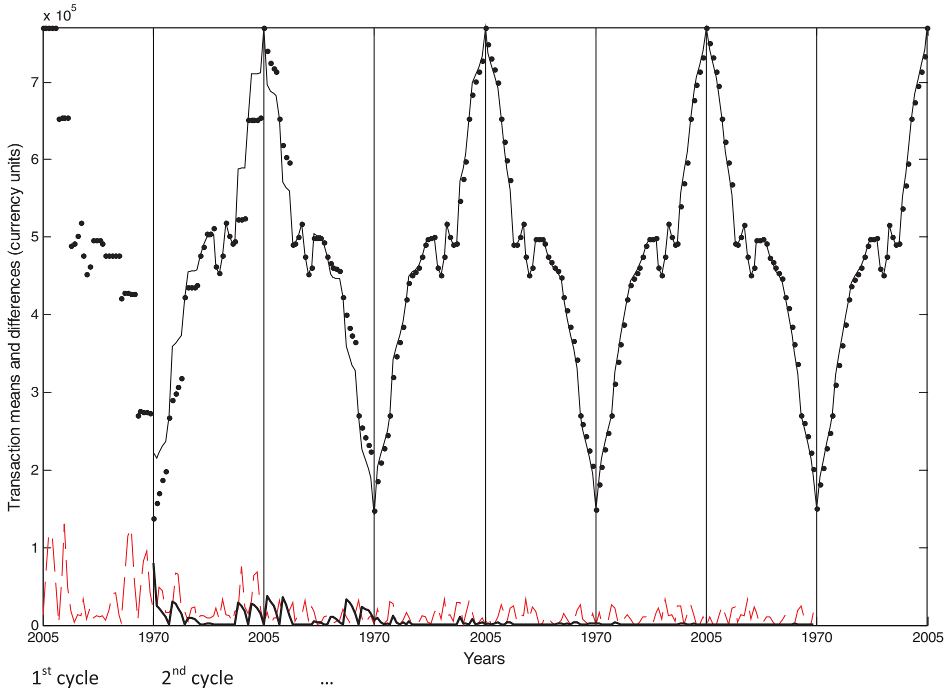


Note: The grey scale units are logarithms of transactions in 2009R$.

according to the sector sequence in $\bar{\mathbf{A}}^{N*}$, and plot the top 5,000 sectors of the re-sorted $\bar{\mathbf{a}}^{N,y*}$ as a topographic map (Figure 6).

First, the accumulation of dark shades towards the top of the figure indicates that in general, sectors that are important in 1 year stay important throughout the time series, as one would expect. Exceptions can be detected as horizontal lines with conspicuous changes of shade between years. Such occurrences may not always mean that a sector has undergone a sudden and substantial expansion or contraction. Instead, such discontinuities could be

FIGURE 7. Analytical measures of fore- and backcast cycling in units of 10^5 2009R$.



Notes: Dotted: transaction means $\tilde{a}^{n,y} = \sum_{i=1}^{I} \frac{a_i^{n,y}}{I}$ of SUTs $\boldsymbol{a}^{n,y}$. Solid thin line: two-cycle moving averages $\bar{\tilde{a}}^{n,y} = 0.5(\tilde{a}^{n,y} + \tilde{a}^{n-1,y})$of transaction means $\tilde{a}^{n,y}$. Dashed line: inter-cycle differences $\tilde{d}^{n,y} = \tilde{a}^{n,y} - \tilde{a}^{n-1,y}$ between transaction means $\tilde{a}^{n,y}$. Solid thick line: inter-cycle differences $\bar{\tilde{d}}^{n,y} = \bar{\tilde{a}}^{n,y} - \bar{\tilde{a}}^{n-1,y}$ between two-cycle moving averages $\bar{\tilde{a}}^{n,y}$

the consequences of imperfect alignment of classifications over time in our concordance matrices (Section 3.1). Thus, a plot such as in Figure 6 is a useful visualisation tool for rapid quality assurance of an entire input–output table time series. Second, during the early 1970s, and between 1997 and 1999, transaction values are markedly lower compared to remaining years, which is a result of discontinuities in the magnitude of gross output (appendix).

In order to demonstrate the effectiveness of fore- and backcast cycling, we plot in Figure 7 first the SUT transaction means $\tilde{a}^{n,y} = \sum_{i=1}^{I} \frac{a_i^{n,y}}{I}$ of tables $\mathbf{a}^{n,y}$ as a function of the cycle-year sequence as in Table 1. Second, we plot the inter-cycle annual difference $\tilde{d}^{n,y} = \tilde{a}^{n,y} - \tilde{a}^{n-1,y}$ between successive transaction means $\tilde{a}^{n-1,y}$ and $\tilde{a}^{n,y}$. For example, $\tilde{d}^{n,y}$ is a measure for the difference between the year-y cycle-n solution $\mathbf{a}^{n,y}$ and the year-y prior-cycle solution $\mathbf{a}^{n-1,y}$. This quantity is useful in showing how with ongoing cycling, the difference between successive SUT iterations $\mathbf{a}^{n-1,y}$ and $\mathbf{a}^{n,y}$ for the same year becomes smaller and smaller as cycling continues. However, it does not become zero because of ongoing hysteresis (compare Figure 2). We therefore compute and plot two-cycle moving averages $\bar{\tilde{a}}^{n,y} = 0.5(\tilde{a}^{n,y} + \tilde{a}^{n-1,y})$, and show that their differences $\bar{\tilde{d}}^{n,y} = \bar{\tilde{a}}^{n,y} + \bar{\tilde{a}}^{n-1,y}$ do become zero as the hysteresis becomes more and more symmetrical (compare Figure 3). The inter-cycle differences between two-cycle moving averages hence provide a good decision aid for how many cycles might suffice in order to arrive at a sufficiently stable final solution for a time series. Note that for this exercise, we purposefully left out the purchasers'-price SUT data (IBGE, 2011), because its inclusion would have led to a continuous series of information post-1990, thus not giving rise to sufficiently obvious data gaps to bridge using the cycling method.

A number of steps are discernible in the transaction means $\tilde{a}^{n,y}$ of the first cycle in Figure 7. These steps mark the transition between years where full input–output information is available and years where only some macroeconomic data such as GDP is available. During the latter years, the input–output structure of the table does not change as significantly, leading to small 'ledges' in the transaction mean curve (dotted). These ledges become drawn out as inter-cycle averaging starts during second cycle.

The differences $\tilde{d}^{n,y}$ between the transaction means $\tilde{a}^{n,y}$ of the first and second cycle are shown by the dashed curve in the first-cycle compartment in Figure 7. The peaks mark transitions that were step-like during the first cycle but drawn out during the second cycle. Note that these inter-cycle differences amount to up to 10^5 2009R$, representing between 15% (2000s) and 30% (pre-1980) of the transaction means themselves. This example demonstrates how fore- and backcasting yield substantially different results, which is perhaps the best justification for why cycling and averaging as described in this article is needed.

As cycling continues, fore- and backcasting trajectories $\tilde{a}^{n,y}$ become more and more similar, and their inter-cycle differences $\tilde{d}^{n,y}$ decrease. Note that during the third cycle, these inter-cycle differences are all below 50,000 2009R$ or below 10% of the transaction means. During the fifth cycle, the differences are all below 30,000 2009R$. Nevertheless, hysteresis continues, and the inter-cycle differences between the transaction means do not disappear even during the eighth cycle.

As illustrated in Figure 2, as hysteresis becomes more and more symmetrical, two-cycle moving averages $\bar{\tilde{a}}^{n,y}$ also become more and more similar, but in contrast to the transaction means $\tilde{a}^{n,y}$, their inter-cycle differences $\bar{\tilde{d}}^{n,y}$ decrease towards zero (compare Figure 3 and thick solid curve in Figure 7). In our example, we can see that four cycles suffice in order to obtain a satisfactory final SUT result from two-cycle moving averages.

5. CONCLUSIONS

Constructing time series of input–output tables has in the past been a labour- and time-intensive undertaking, which partly explains the current lack of comprehensive input–output table time series. Using the example of the Brazilian SUTs from 1970 to 2008, we have presented a new method for constructing time series of tables, featuring a number of innovations: First, we use a software tool equipped with highly automated procedures for assembling initial estimates and constraints, flexible optimisation algorithms for matrix balancing, a powerful graphical user interface, as well as diagnostic and visualisation tools. Our approach is to start with an initial estimate for only 1 base year, which is then balanced in order to obey constraints posed by superior data, and fundamental input–output balance requirements. The base-year solution generated by the optimisation algorithm handling the balancing is then passed on into the subsequent year's balancing procedure as an initial estimate. As such we are able to automatically sweep-optimise an entire time series.

Second, we have developed an effective solution to the problem of integrating data with temporal gaps. In particular, we have dealt with the problem of hysteresis that occurs when input–output transactions are dependent on whether they are the result of forecasting or backcasting over time. The basic idea is to use concurrent sweeps of the time series period, alternating between fore- and backcasting, and taking concurrent averages between fore- and back-cast solutions. We show that as such cycling continues, the time series solution becomes more and more stable.

Whilst repeated cycling effectively irons out any discontinuities and gaps in time series applications, analysts may want to include some sort of inter-year scaling of initial estimates in practice. For example, when years without input–output tables available are constrained by GDP only as in this study, value-added adjusts as a consequence, but intermediate demand is not affected by the optimisation procedure for those years. Therefore, the proportion between the value-added and intermediate-demand blocks may become distorted during the initial sweeps (Wood, personal communication, 16 February 2011). Those distortions would then be almost completely removed during further cycling. Perhaps, improved solutions would be to iteratively pre-scale the initial estimate (Lenzen et al., 2010a) or to start with a set of separate initial estimates constructed via regression (Wood, 2011). We did not discuss such variants in detail in order to clearly distinguish the effects of hysteresis as the main focus of this article.

The method proposed here has clear advantages also over existing parallel methods that could potentially deal with the hysteresis problem, in that our method works even with supporting data that are grossly insufficient (for example, too aggregated or too incomplete) for constructing the required harmonised base- and final-year initial estimates.

We believe our research to be of use for analysts who wish to construct time series of any kind of contingency tables (for example, input–output tables, SUTs, environmental or social satellite accounts, etc.), and who face a limited time and labour budget. Our hope is that with increasing use of automated tools such as the one described in this article, such time series will be constructed more comprehensively, more frequently and more timely in the future.

Acknowledgements

The work described in this article was financially supported by the Australian Research within the scope of Discovery Project DP0985522 and its Linkage Project LP0669290, and by the Research Council of Norway within the scope of Project 30411–196090/V10, administered by the Center for International Climate and Environmental Research in Oslo. The authors are grateful to Professor Joaquim Guilhoto for providing an electronic copy of the 1980 imports table. Richard Wood provided valuable comments that influenced earlier drafts of this article. The authors are also grateful for the scrutiny of the manuscript by an anonymous referee.

References

Alcántara, V. and R. Duarte (2004) Comparison of Energy Intensities in European Union Countries. Results of a Structural Decomposition Analysis. *Energy Policy*, 32, 177–189.

Baiocchi, G. and J. Minx (2010) Understanding Changes in the UK's CO_2 Emissions: A Global Perspective. *Environmental Science & Technology*, 44, 1177–1184.

BBC. (2008) *UK in 'delusion' Over Emissions*. BBC News http://news.bbc.co.uk/today/hi/today/newsid, London, UK, British Broadcasting Corporation.

Carvalheiro, N. (1998) Observações sobre a elaboração da matriz de insumo-produto. *Pesquisa & Debate*, 9, 139–157.

Casler, S.D. and A. Rose (1998) Structural Decomposition Analysis of Changes in Greenhouse Gas Emissions in the US. *Environmental and Resource Economics*, 11, 349–363.

Chang, Y.F. and S.J. Lin (1998) Structural Decomposition of Industrial CO_2 Emissions in Taiwan: An Input–Output Approach. *Energy Policy*, 26, 5–12.

Chen, C.-Y. and A. Rose (1990) A Structural Decomposition Analysis of Changes in Energy Demand in Taiwan: 1971–1984. *Energy Journal*, 11, 127–146.

De Haan, M. (2001) A Structural Decomposition Analysis of Pollution in the Netherlands. *Economic Systems Research*, 13, 181–196.

Han, X. (1995) Structural Change and Labor Requirement of the Japanese Economy. *Economic Systems Research*, 7, 47–65.

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

IBGE. (2008) *Matriz de Insumo-Produto Brasil 2000/2005*. Contas Nacionais número 23, Internet site http://www.ibge.gov.br/home/estatistica/economia/matrizinsumo_produto/publicacao.pdf, Rio de Janeiro, Brazil, Fundação IBGE, Diretoria de Pesquisa, Coordenação de Contas Nacionais.

IBGE. (2010) *Matriz de Insumo-Produto Brasil*. Internet site www.ibge.gov.br/home/estatistica/economia/matrizinsumo_produto/default.shtm, Rio de Janeiro, Brazil, Instituto Brasileiro de Geografia e Estatística, Ministério de Planejamento e Orçamento.

IBGE. (2011) *Sistema de Contas Nacionais – Tabelas Completas*. Internet site http://www.ibge.gov.br/home/estatistica/economia/contasnacionais/2008/defaulttabzip.shtm, Rio de Janeiro, Brazil, Instituto Brasileiro de Geografia e Estatística, Ministério de Planejamento e Orçamento.

IPEA. (2010) *Contas nacionais, Séries históricas*. ipeadata http://www.ipeadata.gov.br/ipeaweb.dll/ipeadata?SessionID=1571094915&Tick=1285649930042&VAR_FUNCAO=Ser_Hist%28127%29&Mod=M, Brasília, Distrito Federal, Brazil, Instituto de Pesquisa Econômica Aplicada.

Jackson, R.W. and A.T. Murray (2004) Alternative Input–Output Matrix Updating Formulations. *Economic Systems Research*, 16, 135–148.

Kagawa, S. and H. Inamura (2001) A Structural Decomposition of Energy Consumption Based on a Rectangular Input–Output Framework: Japan's Case. *Economic Systems Research*, 13, 339–363.

Lahr, M.L. and L. de Mesnard (2004) Biproportional Techniques in Input–Output Analysis: Table Updating and Structural Analysis. *Economic Systems Research*, 16, 115–134.

Lenzen, M., B. Gallego and R. Wood (2006) A Flexible Approach to Matrix Balancing Under Partial Information. *Journal of Applied Input–Output Analysis*, 11&12, 1–24.

Lenzen, M., B. Gallego and R. Wood (2009) Matrix Balancing Under Conflicting Information. *Economic Systems Research*, 21, 23–44.

Lenzen, M., K. Kanemoto, A. Geschke, D. Moran, P.J. Muñoz, J. Ugon, R. Wood and T. Yu (2010a) A Global Multi-Region Input–Output Time Series at High Country and Sector Detail. In: J.M. Rueda-Cantuche and K. Hubacek (eds.) *18th International Input–Output Conference*, http://www.iioa.org/files/conference-1/37_20100617021_Lenen&al_GlobalMRIO_18thIOConf2010.pdf. Sydney, Australia.

Lenzen, M., R. Wood and T. Wiedmann (2010b) Uncertainty Analysis for Multi-Region Input–Output Models – A Case Study of the UK's Carbon Footprint. *Economic Systems Research*, 22, 43–63.

Lin, X. and K.R. Polenske (1995) Input–Output Anatomy of China's Energy Use Changes in the 1980s. *Economic Systems Research*, 7, 67–84.

Los, B. and R. Stehrer (2010) A Preliminary Series of Worldwide Intercountry Input–Output Tables. In: J.M. Rueda-Cantuche and K. Hubacek (eds.) *18th International Input–Output Conference*. Sydney, Australia.

Mukhopadhyay, K. and D. Chakraborty (1999) India's Energy Consumption Changes During 1973/74 to 1991/1992. *Economic Systems Research*, 11, 423–438.

Muñoz, P.J. and K. Hubacek (2008) Material Implication of Chile's Economic Growth: Combining Material Flow Accounting (MFA) and Structural Decomposition Analysis (SDA). *Ecological Economics*, 65, 136–144.

Robinson, S., A. Catteano and M. El-Said (2001) Updating and Estimating a Social Accounting Matrix Using Cross-Entropy Methods. *Economic Systems Research*, 13, 47–64.

Rose, A. (1999) Input–Output Structural Decomposition Analysis of Energy and the Environment. In: J.C.J.M. van den Bergh (ed.) *Handbook of Environmental and Resource Economics*. Cheltenham: Edward Elgar, 1165–1179.

Skolka, J. (1989) Input–Output Structural Decomposition Analysis for Austria. *Journal of Policy Modeling*, 11, 45–66.

Statistics Denmark (2011) *Danish Annual Input–Output Tables*. Internet site http://www.dst.dk/HomeUK/Statistics/ofs/NatAcc/IOTABLES.aspx, Copenhagen, Denmark, Statistics Denmark.

Tarancon, M. and P. Del Rio (2005) Projection of Input–Output Tables by means of Mathematical Programming Based on the Hypothesis of Stable Structural Evolution. *Economic Systems Research*, 17, 1–23.

Temurshoev, U., N. Yamano and C. Webb (2011) Projection of Supply and Use Tables: Methods and Their Empirical Assessment. *Economic Systems Research*, 23, 91–123.

UN. (2010) *CNAE – Classificação Nacional de Atividades Econômicas*. Internet site http://unstats.un.org/unsd/cr/ctryreg/ctrydetail.asp?id=297, United Nations Statistical Division.
Van der Ploeg, F. (1988) Balancing Large Systems of National Accounts. *Computer Science in Economics and Management*, 1, 31–39.
Wachsmann, U., R. Wood, M. Lenzen and R. Schaeffer (2009) Structural Decomposition of Energy Use in Brazil from 1970 to 1996. *Applied Energy*, 86, 578–587.
Wang, E.C. (1996) A Multiplicative Decomposition Method to Identify the Sectoral Changes in Various Developmental Stages: Taiwan, 1966–1991. *Economic Systems Research*, 8, 63–79.
Wiedmann, T., R. Wood, M. Lenzen, J. Minx, D. Guan and J. Barrett (2010) The Carbon Footprint of the UK – Results from a Multi-Region Input–Output model. *Economic Systems Research*, 22, 19–42.
Wier, M. (1998) Sources of Changes in Emissions from Energy: A Structural Decomposition Analysis. *Economic Systems Research*, 10, 99–112.
Wier, M. and B. Hasler (1999) Accounting for Nitrogen in Denmark – A Structural Decomposition Analysis. *Ecological Economics*, 30, 317–331.
Wood, R. (2009) Structural Decomposition Analysis of Australia's Greenhouse Gas Emissions. *Energy Policy*, 37, 4943–4948.
Wood, R. (2011) Construction, Stability and Predictability of an Input–Output Time-Series for Australia. *Economic Systems Research*, 23, 175–211.
Wood, R. and M. Lenzen (2009) Aggregate Measures of Complex Economic Structure and Evolution – A Review and Case Study. *Journal of Industrial Ecology*, 13, 264–283.
Wood, R., M. Lenzen and B. Foran (2009) A Material History of Australia: Evolution of Material Intensity and Drivers of Change. *Journal of Industrial Ecology*, 13, 847–862.
Yamakawa, A. and G. Peters (2011) Structural Decomposition Analysis of greenhouse gas emissions in Norway 1990–2002. *Economic Systems Research*, 23, 303–318.

APPENDIX

There exist major conceptual differences in the purchasers'-price SUT time series (right; IBGE, 2011) and in the basic-price SUT time series (Figure A1; IBGE, 2008; Wachsmann et al., 2009). In the purchasers'-price SUT time series, margins $\mathbf{m}$, net taxes on products $\mathbf{t_p}$, and imports $\mathbf{M}$, are appended as vectors below the supply block $\mathbf{V}$ in order to facilitate the product balance $\mathbf{x_{c^*}}$. This is because in the purchasers'-price SUT time series, margins, net taxes and imports are missing as explicit complements to value added, but rather instead being incorporated into the use ($\mathbf{U}$) and final demand matrices ($\mathbf{y}$). The industry balance $\mathbf{x_i}$ is unaffected by these conceptual differences, though.

As a consequence, use and final demand matrices in the purchasers'-price SUTs cannot be used to inform our basic-price time series. We have hence only used the data items shown shaded in Figure A1.

FIGURE A1. Supply-use system in the purchasers'-price SUT time series (right; IBGE, 2011) and in the basic-price SUT time series (left; IBGE, 2008; Wachsmann et al., 2009).

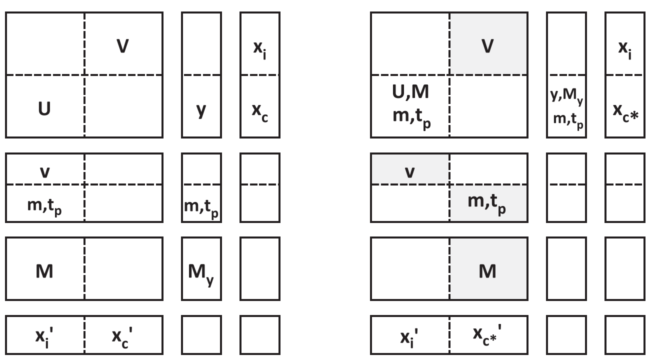
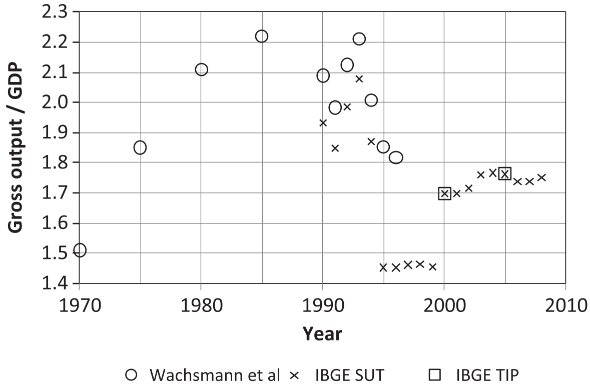


FIGURE A2. Ratio of gross output to GDP in the source data sets from Wachsmann et al. (2009), the official Brazilian SUTs (IBGE, 2011) and the official input–output tables (Tabelas Insumo-Produto, TIP, IBGE, 2010).



In addition, there are a number of additional discontinuities both in the purchasers'-price SUT time series and the basic-price time series used by Wachsmann et al. (2009) (Figure A2). The purchasers'-price SUTs are available in different formats and classifications for the years 1990–1994, 1995–1999 and 2000–2008. As a result, the ratio between gross output and GDP, and hence the ratio between data components such as the supply matrix and the value-added block, are subject to discontinuities. For the three blocks listed above, these ratios are 2.0, 1.5 and 1.75, respectively (see $\times$symbols).

These discontinuities have consequences for matrix balancing, as well as for analytical results obtained through, amongst other techniques, the classical Leontief demand-pull calculus or SDA. The system feedback inherent in intermediate demand is much lower for the years 1995–1999 compared to the remaining periods, and hence multipliers and SDA terms can be expected to be lower for these years.

Similar differences exist within the time series constructed by Wachsmann et al. (2009): the 1970, 1975 and 1980 SUTs are constructed from raw data in completely different classifications, and at least the gross-output-to-GDP ratios are quite different for 1970 and 1975.