

Accepted Manuscript

Hybrid life cycle assessment (LCA) does not necessarily yield more accurate results than process-based LCA

Yi Yang, Reinout Heijungs, Miguel Brandão



PII: S0959-6526(17)30440-7

DOI: [10.1016/j.jclepro.2017.03.006](https://doi.org/10.1016/j.jclepro.2017.03.006)

Reference: JCLP 9133

To appear in: *Journal of Cleaner Production*

Received Date: 13 December 2016

Revised Date: 4 February 2017

Accepted Date: 1 March 2017

Please cite this article as: Yang Y, Heijungs R, Brandão M, Hybrid life cycle assessment (LCA) does not necessarily yield more accurate results than process-based LCA, *Journal of Cleaner Production* (2017), doi: 10.1016/j.jclepro.2017.03.006.

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

[word count: 3950]

Hybrid life cycle assessment (LCA) does not necessarily yield more accurate results than process-based LCA

Yi Yang^{1*}, Reinout Heijungs^{2,3}, and Miguel Brandão^{4,5}

1. CSRA Inc., USA

2. Department of Econometrics and Operations Research, Vrije Universiteit Amsterdam, Netherlands

3. Department of Industrial Ecology, Institute of Environmental Sciences, Leiden University, Netherlands

4. Department of Sustainable Development, Environmental Science and Engineering, KTH - Royal Institute of Technology, Sweden

5. Department of Bioeconomy and Systems Analysis, Institute of Soil Science and Plant Cultivation, Poland

Abstract

Hybrid life cycle assessment (LCA), through combining input-output (IO) models and process-based LCA for a complete system boundary, is widely accepted by its practitioners as a more accurate approach than process-based LCA with an incomplete system boundary. Without a complete process model for verification, however, the performance of hybrid LCA remains unclear. Here, using a counterexample we show that hybrid LCA does not necessarily provide more accurate results than process-based LCA, simply because the aggregation of heterogeneous processes in IO models may introduce more errors. In so doing, we demonstrate that only when IO-based LCA and process-based LCA have the same level of detail would they yield the same results. Whether hybrid LCA provides more accurate estimates depends on whether the IO model introduced serves as an adequate proxy for the missing products as opposed to if they were estimated by a complete process model. The use of a highly-aggregated IO model runs the risk of overestimation, and could result in a larger relative error than the truncation error resulting from an incomplete process model. Our study seeks to provide a balanced view of hybrid LCA, and our findings offer important insights for future hybrid LCA studies to improve the accuracy and realm of applicability of the approach.

Key words

life cycle assessment; process; input-output; hybrid; system boundary; aggregation; truncation error;

1. Introduction

Life cycle assessment (LCA) has traditionally used linear models to address various questions related to product environmental sustainability (Yang, 2016). Under the linear framework for Life Cycle Inventory analysis (LCI), three computational methods can be distinguished, process, Input-Output (IO), and hybrid (Heijungs and Suh, 2002). Process-based LCA uses detailed process-level information reflecting how the unit processes of a product system are interconnected through commodity flows (Suh and Huppes, 2005). IO-based LCA relies on aggregate sectoral information from IO tables that record annual transactions between all the productive activities of an economy. A typical IO table has dozens to hundreds of sectors (Miller and Blair, 2009). Hybrid LCA is a combination of process- and IO-based LCA, with different degrees of integration between the two constituent methods (Suh and Huppes, 2005).

Hybrid LCA has been developed as a compromise to correct for the truncation error of process LCA-based while taking advantage of its process specificity. The rationale is as follows. Although process-based LCA provides detailed information, it suffers from system incompleteness. This is because, so the argument goes, all processes of an economy are connected with one another directly or indirectly (Suh et al., 2004), thus the LCI of a, or any, product produced by an economy should in theory cover all the processes involved. Given the lack of such a complete process database, process-based LCA in practice covers only a subset of an economy and thus leads to a truncation error in estimating life cycle emissions, resource extractions, and impacts. The size of the truncation error could be up to 50% (Lenzen, 2001). By contrast, IO-based LCA is complete as far as productive sectors are concerned, but it is too aggregate; for example, different types of plastic are merged into an overall plastic sector. Through combining process-based LCA in the detailed foreground system and IO-based LCA as a more generic background system, hybrid LCA is argued to achieve both specificity and system completeness (Suh et al., 2004). Because of this merit, hybrid LCA is sometimes considered a more advanced method than process- or IO-based LCA (Minx et al., 2008; Suh et al., 2004; Suh and Huppes, 2005), representing “the current state-of-the art in ecological economic modelling” (Wiedmann and Minx, 2008).

Despite the recognition of hybrid LCA as an advanced method and its increasing number of applications (Nakamura and Nansai, 2016), its performance in practice remains unclear given the lack of a complete process LCA database. Clearly, hybrid LCA would yield larger estimates than process-based LCA simply because of the addition of IO models to make up for missing parts from the foreground process system. However, larger estimates do not necessarily imply better estimates, in the sense of being closer to the truth, where truth is conceived as derived from a complete process database. The questions we seek to

address in this paper are: is hybrid LCA more accurate than process-based LCA? If so, under what circumstances?

In the following section, we demonstrate using an example that theoretically, hybrid LCA is not necessarily more accurate than process-based LCA simply due to aggregation errors introduced on the part of IO. Our study contributes to the literature by providing a more balanced view of hybrid LCA and calling attention to possibly uncritical application of the approach. Our findings on aggregation errors of IO models can be investigated further to improve the accuracy of hybrid LCA, and to develop criteria for the realm of applicability of hybrid LCA. We should not default to hybridization with IO models without critically evaluating if such efforts would clearly lead to improved estimates.

2. A simple hypothetical example

In the absence of a complete process LCA database that covers all components of an economy, how do we investigate if hybrid LCA delivers better results than process-based LCA? We address this question from an angle similar to that of Lenzen's in his seminal work on the potential truncation error of process-based LCA (Lenzen, 2001). Suppose we have a complete process database that covers every single process/product of an economy, in which case we can calculate the "true" LCA results of any product. We can then take part of the database out and make an incomplete process-based LCA model. Suppose we also have price and output information for all the products covered in the complete database, we can then make an IO-based LCA model by aggregating certain processes. And we can further make use of the process- and IO-based LCA models and derive a hybrid model. The "truth" derived from the complete process-based LCA model will serve as a benchmark against which we can measure how accurate the incomplete process model, the aggregated IO model, and the hybrid LCA model are.

We can, for example, construct a 100,000 process database, or even larger, with fictitious numbers that make economic and technological sense, and put the above ideas to test. But in fact, a much smaller system is already able to demonstrate the core of the argument above because the principles are the same. Below we construct a 5-process example to demonstrate that there is no evidence that hybrid LCA, argued to have a more complete system boundary, is more accurate than process-based LCA with an incomplete system boundary. The key to the accuracy of hybrid LCA is whether the IO model introduced can adequately represent the missing products as opposed to if they were represented by a complete process model. Because of the aggregation errors on the part of IO, hybrid LCA can lead to a relative error larger than that of incomplete process-based LCA. Only when an IO model has the same level of detail as a process model would it yield the same results.

2.1. Case 0: A complete process-based system

Suppose an economy consists of five processes, namely, corn cultivation, wheat cultivation, energy generation, machine production, and popcorn making. Each process produces one unique product, and they all generate CO₂ emissions only. The technological and environmental aspects of the processes are shown below in matrix formulation (Heijungs and Suh, 2002):

$$\mathbf{A}_0 = \begin{array}{ccccc} & \begin{array}{c} \text{corn cultivation} \\ \text{wheat cultivation} \\ \text{energy generation} \\ \text{machine production} \\ \text{popcorn making} \end{array} & & & \\ \begin{array}{c} \text{corn (kg)} \\ \text{wheat (kg)} \\ \text{energy (MJ)} \\ \text{machine (unit)} \\ \text{popcorn (kg)} \end{array} & \begin{bmatrix} 1 & 0 & 0 & 0 & -0.5 \\ 0 & 1 & -0.02 & 0 & 0 \\ -0.2 & -0.1 & 1 & -0.3 & -0.7 \\ -0.01 & -0.02 & -0.04 & 1 & -0.03 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} & & & \end{array} \quad (1)$$

$$\mathbf{B}_0 = [2 \quad 7 \quad 1.5 \quad 3 \quad 1] \text{ CO}_2 \text{ (kg)} \quad (2)$$

where \mathbf{A}_0 is the technology matrix and \mathbf{B}_0 the environmental matrix. Columns in \mathbf{A}_0 and \mathbf{B}_0 indicate processes; rows in \mathbf{A}_0 indicate products with negative signs representing use and positive signs production; and rows in \mathbf{B}_0 indicate emissions or resource use. To produce 1 kg of corn, for example, requires 0.2 MJ of energy and 0.01 unit of machine while generating 2 kg of CO₂.

Suppose our functional unit is 1 kg of popcorn,

$$\mathbf{f}_0 = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} \quad (3)$$

then its cradle-to-gate CO₂ emissions are:

$$\mathbf{m}_0 = \mathbf{B}_0 \mathbf{A}_0^{-1} \mathbf{f}_0 = [3.55] \quad (4)$$

The result shows that the popcorn's aggregated CO₂ emissions are 3.55 kg/kg. Although popcorn production releases only 1 kg CO₂ directly (per kg), its upstream emissions are over twice as much.

Case 0 is our benchmark for all other analyses below. We shall regard it as the “true” situation. In any of the analyses below, it will be assumed that the data of case 0 are known.

2.2. Case 1: An incomplete process-based system

Now let us suppose that the corn process from the database above is unknown or missing, and we therefore have a process model with an incomplete system boundary, as shown below:

$$\mathbf{A}_1 = \begin{array}{c} \begin{array}{cccc} & \text{wheat cultivation} & \text{energy generation} & \text{machine production} & \text{popcorn making} \\ \begin{bmatrix} 1 & -.02 & 0 & 0 \\ -.1 & 1 & -.3 & -.7 \\ -.02 & -.04 & 1 & -.03 \\ 0 & 0 & 0 & 1 \end{bmatrix} & \begin{array}{l} \text{wheat (kg)} \\ \text{energy (MJ)} \\ \text{machine (unit)} \\ \text{popcorn (kg)} \end{array} \end{array} \end{array} \quad (5)$$

$$\mathbf{B}_1 = [7 \quad 1.5 \quad 3 \quad 1] \text{ CO}_2 \text{ (kg)} \quad (6)$$

Again, we are to calculate the aggregated CO₂ emissions of 1 kg of popcorn,

$$\mathbf{f}_1 = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} \quad (7)$$

using the standard formula:

$$\mathbf{m}_1 = \mathbf{B}_1 \mathbf{A}_1^{-1} \mathbf{f}_1 = [2.36] \quad (8)$$

Compare equations (8) and (4): clearly, the absence of corn, reflecting a missing part of the system, leads to an underestimate of the CO₂ emissions for popcorn, in this example with a relative error of 33.7%.

Case 1 represents the typical situation in process-based LCA. It has reasonably detailed process information, but is incomplete, leading to underestimates of life cycle environmental impact.

2.3. Case 0 revisited: An unaggregated IO-based system

We return to case 0. Suppose we have further information on product quantities (\mathbf{q}_0) and unit prices (\mathbf{p}_0) for the five products:

$$\mathbf{q}_0 = \begin{bmatrix} 50 \\ 100 \\ 200 \\ 20 \\ 40 \end{bmatrix} \begin{matrix} \text{corn (kg)} \\ \text{wheat (kg)} \\ \text{energy (MJ)} \\ \text{machine (unit)} \\ \text{popcorn (kg)} \end{matrix} \quad \mathbf{p}_0 = \begin{bmatrix} 2 \\ 1.4 \\ 1 \\ 10 \\ 4 \end{bmatrix} \begin{matrix} \text{corn (\$/kg)} \\ \text{wheat (\$/kg)} \\ \text{energy (\$/MJ)} \\ \text{machine (\$/unit)} \\ \text{popcorn (\$/kg)} \end{matrix} \quad (9)$$

with which we can calculate the value of the total output of each product (\mathbf{x}_0) and transform the process database into an IO transaction table (\mathbf{Z}_0) (Miller and Blair, 2009):

$$\mathbf{x}_0 = \widehat{\mathbf{p}}_0 \mathbf{q}_0 = \begin{bmatrix} 100 \\ 140 \\ 200 \\ 200 \\ 160 \end{bmatrix} \quad (10)$$

$$\mathbf{Z}_0 = \widehat{\mathbf{p}}_0 (\mathbf{I} - \mathbf{A}_0) \widehat{\mathbf{q}}_0 = \begin{matrix} & \begin{matrix} \text{corn} & \text{wheat} & \text{energy} & \text{machine} & \text{popcorn} \end{matrix} \\ \begin{bmatrix} 0 & 0 & 0 & 0 & 40 \\ 0 & 0 & 5.6 & 0 & 0 \\ 10 & 10 & 0 & 6 & 28 \\ 5 & 20 & 80 & 0 & 12 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} & \begin{matrix} \text{corn (\$)} \\ \text{wheat (\$)} \\ \text{energy (\$)} \\ \text{machine (\$)} \\ \text{popcorn (\$)} \end{matrix} \end{matrix} \quad (11)$$

where \wedge symbolizes diagonalization, and \mathbf{I} is the identity matrix. What \mathbf{Z}_0 (in combination with \mathbf{x}_0) means is that, for example, to produce \$140 worth's wheat requires \$10 worth of energy and \$20 worth of machine.

We can continue to construct the technology and environmental matrices for now the 5-sector IO-based LCA model:

$$\mathbf{A}_0^* = \mathbf{Z}_0 \widehat{\mathbf{x}}_0^{-1} = \begin{bmatrix} 0 & 0 & 0 & 0 & .25 \\ 0 & 0 & .028 & 0 & 0 \\ .1 & .071 & 0 & .03 & .175 \\ .05 & .143 & .4 & 0 & .075 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \quad (12)$$

$$\mathbf{B}_0^* = \mathbf{B}_0 \widehat{\mathbf{p}}_0^{-1} = [1 \quad 5 \quad 2 \quad .3 \quad .25] \quad (13)$$

where * is used to differentiate the 5-sector IO model from the 5-process model (equations (1-2)). What \mathbf{A}_0^* and \mathbf{B}_0^* mean is that, for example, to produce \$1 worth's energy requires \$0.028 worth of wheat and \$0.4 worth of machine, and meanwhile it generates 2 kg of CO₂ emissions.

Note that the IO-LCA model we have constructed is at the same level of detail as the complete process database (equations 1-2). This is not the case for a “normal” IO model, which is highly aggregated with commonly dozens to hundreds of products lumped together in one sector. This “unnatural” IO-LCA model, however, serves to make a point soon to be clear.

As with previous cases, we calculate the cradle-to-gate CO₂ emissions of 1 unit of popcorn:

$$\mathbf{f}_0^* = \widehat{\mathbf{p}}_0 \mathbf{f}_0 = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 4 \end{bmatrix} \quad (14)$$

by means of

$$\mathbf{m}_0^* = \mathbf{B}_0^*(\mathbf{I} - \mathbf{A}_0^*)^{-1} \mathbf{f}_0^* = [3.55] \quad (15)$$

In this case, we have reformulated the functional unit from 1 unit of popcorn to \$4 given its unit price being \$4. Compare equations (15) and (4): the two models generate the same results. This is not a coincidence and can be proved mathematically: when process-based LCA and IO-based LCA have the same level of detail, they yield the same results (see Appendix).

2.4. Case 2: An aggregated IO-based system

To study the effect of aggregation in IO-based LCA, we aggregate two sectors, for example, corn and wheat into an agriculture sector. With this change, the IO-table becomes more realistic, reflecting the contrast between more detailed process-based information and more aggregated IO-based tables:

$$\mathbf{Z}_2 = \mathbf{iZ}_0\mathbf{j} = \begin{bmatrix} 0 & 5.6 & 0 & 40 \\ 20 & 0 & 6 & 28 \\ 25 & 80 & 0 & 12 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{matrix} \text{agriculture} \\ \text{energy} \\ \text{machine} \\ \text{popcorn} \end{matrix} \begin{matrix} \text{agriculture (\$)} \\ \text{energy (\$)} \\ \text{machine (\$)} \\ \text{popcorn (\$)} \end{matrix} \quad (16)$$

$$\mathbf{i} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (17)$$

$$\mathbf{j} = \mathbf{i}' = \begin{bmatrix} 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \quad (18)$$

$$\mathbf{x}_2 = \mathbf{i}\mathbf{x}_0 = \begin{bmatrix} 240 \\ 200 \\ 200 \\ 160 \end{bmatrix} \quad (19)$$

$$\mathbf{A}_2 = \mathbf{Z}_2\widehat{\mathbf{x}}_2^{-1} = \begin{bmatrix} 0 & .028 & 0 & .25 \\ .083 & 0 & .03 & .175 \\ .104 & .4 & 0 & .075 \\ 0 & 0 & 0 & 0 \end{bmatrix} \quad (20)$$

$$\mathbf{B}_2 = \mathbf{B}_0^*\widehat{\mathbf{x}}_0\widehat{\mathbf{x}}_2^{-1} = [3.3 \quad 2 \quad .3 \quad .25] \quad (21)$$

So the aggregate CO₂ emissions of 1 kg of popcorn are thus:

$$\mathbf{f}_2 = \mathbf{i}\mathbf{f}_0^* = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 4 \end{bmatrix} \quad (22)$$

$$\mathbf{m}_2 = \mathbf{B}_2(\mathbf{I} - \mathbf{A}_2)^{-1}\mathbf{f}_2 = [5.84] \quad (23)$$

Compare equation (23) with (4) or (15); the estimate given by the aggregate IO-based LCA model is considerably larger, with a relative error of 64.3%.

2.5. Case 3: A hybrid system

Let us move on to building a hybrid LCA model using the incomplete process model and the aggregated IO model. Three types of hybrid model have been differentiated: tiered, IO-based, and integrated (Suh and Huppes, 2005). The tiered analysis, which uses an IO-model to fill in the gaps created by a process-based model, seems to be the most widely applied hybrid approach (Nakamura and Nansai, 2016). Therefore, we focus on tiered hybrid model here, shown as below.

$$\mathbf{A}_3 = \begin{bmatrix} \mathbf{A}_1 & \mathbf{0} \\ \mathbf{U} & \mathbf{I} - \mathbf{A}_2 \end{bmatrix} = \begin{bmatrix} \text{wheat cultivation} & \text{energy generation} & \text{machine production} & \text{popcorn making} & \text{agriculture} & \text{energy} & \text{machine} & \text{popcorn} \\ \begin{bmatrix} 1 & -0.02 & 0 & 0 & 0 & 0 & 0 & 0 \\ -0.1 & 1 & -0.3 & -0.7 & 0 & 0 & 0 & 0 \\ -0.02 & -0.04 & 1 & -0.03 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 & 1 & -0.028 & 0 & -0.25 \\ 0 & 0 & 0 & 0 & -0.083 & 1 & -0.03 & -0.175 \\ 0 & 0 & 0 & 0 & -0.104 & -0.4 & 1 & -0.075 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} & \begin{bmatrix} \text{wheat (kg)} \\ \text{energy (MJ)} \\ \text{machine (unit)} \\ \text{popcorn (kg)} \\ \text{agriculture (\$)} \\ \text{energy (\$)} \\ \text{machine (\$)} \\ \text{popcorn (\$)} \end{bmatrix} \end{bmatrix} \quad (24)$$

where \mathbf{U} is the upstream matrix that connects the IO-table with the missing part of the process matrix:

$$\mathbf{U} = \begin{bmatrix} 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \quad (25)$$

The value -1 is a result of the input of 0.5 kg of corn, which is missing in the incomplete process system \mathbf{A}_1 (see equation (1)), multiplied by its price, 2 \$/kg. And it appears in the same row as the IO sector (agriculture), indicating that an input of \$1 worth of agriculture to produce 1 kg of popcorn. In other words, the aggregated corn and wheat sector from the IO model is used as a proxy for the missing corn product in the incomplete process model. The environmental matrix is given by

$$\mathbf{B}_3 = [\mathbf{B}_1 \quad \mathbf{B}_2] = [7 \quad 1.5 \quad 3 \quad 1 \quad 3.3 \quad 1.5 \quad .3 \quad .25] \quad (26)$$

Now we can calculate the cradle-to-gate CO₂ emissions per kg of popcorn using the tiered hybrid approach:

$$\mathbf{f}_3 = \begin{bmatrix} \mathbf{f}_1 \\ \mathbf{0} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} \quad (27)$$

The emissions are found as follows:

$$\mathbf{m}_3 = \mathbf{B}_3 \mathbf{A}_3^{-1} \mathbf{f}_3 = [5.87] \quad (28)$$

Compare equation (28) with (8) and (4) or (15): while the estimate given by the tiered hybrid model is understandably larger than that given by the incomplete process model, it is also larger than the “truth.” The relative error of the tiered hybrid model is 65.3%, which is larger than the error (33.7%) of the incomplete process-based model. In this example, therefore, hybrid LCA does not outperform process-based LCA with an incomplete system boundary.

2.6. The accuracy of hybrid LCA

In the example above, although hybrid (or IO-based) LCA covers a complete system boundary for popcorn, it also yields a much larger estimate, with a larger relative error as compared with results of the complete process model. This is because the aggregation of corn and wheat in the IO model has resulted in a much higher emission rate for corn as opposed to that given by the perfect process-based model or unaggregated IO-based model. The embodied emissions of corn (kg/kg) given by the complete process-based model are:

$$\mathbf{m}_0 = \mathbf{B}_0 \mathbf{A}_0^{-1} \mathbf{f}_0 = [2.39] \quad (30)$$

whereas in the tiered hybrid model, they are, as calculated by the IO-based LCA model:

$$\mathbf{m}_3 = \mathbf{B}_3 \mathbf{A}_3^{-1} \mathbf{f}_3 = [7.03] \quad (31)$$

This large aggregation error results from the fact that the two products merged, corn and wheat, are quite different, and particularly that the emission intensity of wheat is much higher than that of corn (7 kg/kg or 5 kg/\$ versus 2 kg/kg or 1 kg/\$). If we change the emission intensity of wheat to 2 kg/kg in equation (2), making the aggregated IO sector more homogenous, the error of hybrid LCA becomes smaller than that of the incomplete process model. This is because now the embodied emission rates for corn as given by the complete process model and IO model are much closer (2.37 kg/kg versus 2.85 kg/kg). In this case, the introduction of an IO model is beneficial, and hybrid LCA yields more accurate results than incomplete process-based LCA. The interested reader can play with the example to gain a better understanding (see the supporting information (SI)).

3. Discussion and conclusions

The example above shows that due to sectoral aggregation, hybrid (or IO-based) LCA, although covering a complete system boundary, is not necessarily more accurate than a process model with an incomplete system boundary. One may argue that tiered hybrid runs the risk of double counting (Strømman et al., 2009) and perhaps IO-based or integrated hybrid LCA (Nakamura and Nansai, 2016) would perform better than process LCA. This argument, however, misses the real origin of the problem with IO-based or hybrid LCA, namely, sectoral aggregation of products with heterogeneous technological and environmental profiles. Because products aggregated into a sector can be substantially different, some of them are bound to be significantly misrepresented by sectoral averages. To further illustrate the problem with sectoral aggregation, we constructed another example showing that IO-based hybrid LCA, which avoids the issue of double counting, does not necessarily generate more accurate results than process-based LCA (see SI).

IO-based LCA may be adequate for questions at the macro or meso level, but it may fall short of product-level analyses, given the aggregation errors shown above. For hybrid LCA, whether the addition of an IO model would improve the life cycle estimates depends on whether the missing inputs from the process model can be adequately represented by the IO model, as opposed to if they were estimated by a complete process model. As the example above reflects, this further depends on the magnitude of the aggregation errors in the IO model. If the missing products are represented by IO sectors that are aggregates of widely different processes/products, hybrid LCA runs the risk of overestimation (equations (28)). Consequently, it may result in an even larger relative error than the truncation error from an incomplete process model. For example, if we are missing cabbage in our process model, it is unclear whether we will improve the life-cycle results by filling the gap with the agriculture sector from a highly aggregated IO model such as that of China (Yang and Suh, 2011), or even with the vegetable sector from a much more disaggregated

IO model such as that of the US (Weber et al., 2009). This is because the environmental profiles of agricultural commodities are so different (Yang et al., 2016) that an aggregated category (e.g., grains and vegetables) is unrepresentative of the majority of the commodities aggregated, and may do more harm than provide cures.

On the other hand, if the added IO model yields similar or smaller estimates for the missing products compared to that estimated by a complete process model, hybrid LCA would be more accurate than an incomplete process LCA. Especially, if the IO sectors are highly specific or their aggregated products are relatively homogenous, hybrid LCA may provide more accurate results than process LCA. For example, if the missing product is electricity, hybridization may be beneficial as electricity generation is in general adequately represented in IO models partly for being an individual sector unmerged with other processes.

A further compounding issue is how we bridge process- and IO-based models (e.g., equation (25)). The example constructed above assumes perfect information on product prices and consistent data years. In practice, however, price information for certain products may not be readily available for reasons such as being proprietary, strongly fluctuating, or distorted through tariffs or subsidies. And data years may be very different for the process model and the IO model. These problems often result in the reliance on assumptions and expert judgements that are rarely verified in the process of hybridization.

Given these arguments, caution should be exercised when hybridizing process- and IO-based LCA models. An uncritical application can render it meaningless, and there are cases in which a process-based model with incomplete system boundary may perform better. In other words, we should not default to hybrid LCA. Instead, we need to critically evaluate if the effort of hybridization would clearly lead to improved estimates. The findings of our study on aggregation errors can be investigated further to improve the accuracy of hybrid LCA, and to develop criteria for the realm of applicability of hybrid LCA. For example, future studies may evaluate the homogeneity or heterogeneity of different IO sectors, to gain a better understanding of whether they would form an adequate basis for hybridization.

Appendix

Here we prove that when IO-based LCA is unaggregated or has the same level of disaggregation as process-based LCA, it yields the same results as the latter, assuming a process produces a unique product. Mathematically, the two models are two sides of the same coin.

In process-based LCA, emissions are calculated by

$$\mathbf{m} = \mathbf{B}\mathbf{A}^{-1}\mathbf{f}$$

With additional information on annual production quantities, \mathbf{q} , as produced by the entire economy defined by \mathbf{A} , and unit prices, \mathbf{p} , we can construct an IO-table that represents this economy. Starting with the sectoral transaction table,

$$\mathbf{Z} = \hat{\mathbf{p}}(\mathbf{I} - \mathbf{A})\hat{\mathbf{q}}$$

where \mathbf{I} is the identity matrix and $\hat{\cdot}$ indicates diagonalization, and together with sector output,

$$\mathbf{x} = \hat{\mathbf{p}}\mathbf{q}$$

we arrive at the technology matrix (i.e., direct requirement coefficients) and environmental matrix of the IO-based LCA model,

$$\mathbf{A}^* = \mathbf{Z}\hat{\mathbf{x}}^{-1}$$

$$\mathbf{B}^* = \mathbf{B}\hat{\mathbf{p}}^{-1}$$

where $*$ differentiates IO from process-based. We assume here that both \mathbf{x} and \mathbf{p} contain no zero elements, so that the expressions $\hat{\mathbf{x}}^{-1}$ and $\hat{\mathbf{p}}^{-1}$ exist. This implies that all sectors produce a non-zero output, and that all products have a non-zero price. Note that the IO-based model operates on monetary terms, so to represent the same functional unit in a Leontief multiplier framework, we define a shock

$$\mathbf{f}^* = \hat{\mathbf{p}}\mathbf{f}$$

Now we are ready to prove that the two models yield the same results.

$$\mathbf{m}^* = \mathbf{B}^*(\mathbf{I} - \mathbf{A}^*)^{-1}\mathbf{f}^* = \mathbf{B}\hat{\mathbf{p}}^{-1}(\mathbf{I} - \mathbf{Z}\hat{\mathbf{x}}^{-1})^{-1}\hat{\mathbf{p}}\mathbf{f} = \mathbf{B}\hat{\mathbf{p}}^{-1}(\mathbf{I} - \mathbf{Z}\hat{\mathbf{p}}\mathbf{q}^{-1})^{-1}\hat{\mathbf{p}}\mathbf{f}$$

Because for any square matrix \mathbf{X} , it is true that $\mathbf{X}\hat{\mathbf{y}} = \hat{\mathbf{y}}\mathbf{X}$, this reduces to

$$\mathbf{m}^* = \mathbf{B}(\mathbf{I} - \mathbf{Z}\hat{\mathbf{p}}\mathbf{q}^{-1})^{-1}\mathbf{f}$$

Finally, we use $\mathbf{Z} = \hat{\mathbf{p}}(\mathbf{I} - \mathbf{A})\hat{\mathbf{q}}$, so that $\hat{\mathbf{p}}^{-1}\mathbf{Z}\hat{\mathbf{q}}^{-1} = \hat{\mathbf{p}}^{-1}\hat{\mathbf{p}}^{-1}\hat{\mathbf{q}}^{-1} = \mathbf{Z}\hat{\mathbf{p}}\mathbf{q}^{-1} = \mathbf{I} - \mathbf{A}$ and therefore

$$\mathbf{m}^* = \mathbf{B}(\mathbf{I} - \mathbf{Z}\hat{\mathbf{p}}\mathbf{q}^{-1})^{-1}\mathbf{f} = \mathbf{B}\mathbf{A}^{-1}\mathbf{f} = \mathbf{m}$$

So in conclusion

$$\mathbf{m}^* = \mathbf{m}$$

which concludes the proof.

References

- Heijungs, R., Suh, S., 2002. The computational structure of life cycle assessment. Kluwer Academic Pub, Dordrecht.
- Lenzen, M., 2001. Errors in Conventional and Input Output—based Life—Cycle Inventories. *J. Ind. Ecol.* 4, 127–148.
- Miller, R., Blair, P., 2009. Input-output analysis: foundations and extensions. Cambridge University Press, Cambridge, England.
- Minx, J.C., Wiedmann, T., Barrett, J., 2008. Methods Review to Support the PAS Process for the Calculation of the Greenhouse Gas Emissions Embodied in Good und Services.
- Nakamura, S., Nansai, K., 2016. Input–Output and Hybrid LCA, in: Finkbeiner, M. (Ed.), *Special Types of Life Cycle Assessment*. Springer.
- Strømman, A.H., Peters, G.P., Hertwich, E.G., 2009. Approaches to correct for double counting in tiered hybrid life cycle inventories. *J. Clean. Prod.* 17, 248–254.
- Suh, S., Huppes, G., 2005. Methods for life cycle inventory of a product. *J. Clean. Prod.* 13, 687–697.
- Suh, S., Lenzen, M., Treloar, G., Hondo, H., Horvath, A., Huppes, G., Jolliet, O., Klann, U., Krewitt, W., Moriguchi, Y., 2004. System boundary selection in life-cycle inventories using hybrid approaches. *Env. Sci Technol* 38, 657–664.
- Weber, C., Matthews, D., Venkatesh, A., Costello, C., H. Scott, M., 2009. The 2002 US benchmark version of the economic input-output life cycle assessment (EIO-LCA) model. Green Design Institute, Carnegie Mellon University, Pittsburgh, PA.
- Wiedmann, T., Minx, J., 2008. A definition of “carbon footprint,” in: *Ecological Economics Research Trends*. NOva Science Publishers, Hauppauge NY, USA, pp. 1–11.
- Yang, Y., 2016. Two sides of the same coin: consequential life cycle assessment based on the attributional framework. *J. Clean. Prod.* 127, 274–281. doi:10.1016/j.jclepro.2016.03.089
- Yang, Y., Suh, S., 2011. Environmental Impacts of Products in China. *Environ. Sci. Technol.* 45, 4102–4109.
- Yang, Y., Tao, M., Suh, S., 2016. Geographic variability of life cycle inventories of major crops in the U.S. *Int. J. Life Cycle Assess.* In review.