EPA Public Access

Author manuscript

Environ Model Softw. Author manuscript; available in PMC 2019 January 01.

About author manuscripts

Submit a manuscript

Published in final edited form as:

Environ Model Softw. 2018 January 1; 99: 52–57. doi:10.1016/j.envsoft.2017.09.017.

Exploring the relevance of spatial scale to life cycle inventory results using environmentally-extended input-output models of the United States

Yi Yang¹, Wesley W. Ingwersen², and David E. Meyer²

¹CSRA Inc., Falls Church, VA 22042

²U.S. Environmental Protection Agency, National Risk Management Research Laboratory, Cincinnati, OH 45268

Abstract

The accuracy of direct and indirect resource use and emissions of products as quantified in life cycle models depends in part upon the geographical and technological representativeness of the production models. Production conditions vary not just between nations, but also within national boundaries. Understanding the level of geographic resolution within large industrial nations needed to reach acceptable accuracy has not been well-tested across the broad spectrum of goods and services consumed. Using an aggregate 15-industry environmentally-extended input-output model of the US along with detailed interstate commodity flow data, we test the accuracy of regionalizing the national model into two-regions (state - rest of US) versus 51 regions (all US states + DC). Our findings show the two-region form predicts life cycle emissions and resources used within 10–20% of the more detailed 51-region form for most of the environmental flows studied. The two-region form is less accurate when higher variability exists in production conditions for a product.

Keywords

spatial scale; life cycle assessment; environmentally-extended input-output; aggregation; regional variation

1. Introduction

Life cycle assessment (LCA) is an established and internationally standardized framework for estimating environmental impacts of goods and services and policies affecting their production, distribution, use, and disposal (ISO, 2006). A widely applied LCA method is the environmentally-extended input-output (EEIO) models, which use sector-level economic statistics in combination with various environmental data processed to represent sector emissions and resource use. The method has known limitations such as the lack of product-

Disclaimer

The views expressed in this article are those of the authors and do not necessarily represent the views or policies of the U.S. Environmental Protection Agency.

specific data, or aggregation, and uncertainties regarding the price-quantity relationships (Heijungs and Suh, 2002). It also has many advantages, including its comprehensiveness and use of data curated by national statistical agencies such as those data collected in a census, or otherwise reported by legal mandate that are considered highly reliable.

EEIO models have been applied at national and global scales to quantify emissions associated with consumption and embodied in trade. A number of national and global EEIO models have been developed (Kerkhof et al., 2009; Lenzen, 1998; McGregor et al., 2008; Nansai, 2009; Weber et al., 2009; Wiedmann, 2009; Yang and Suh, 2011), including the recent USEEIO model created by the authors and others (Yang et al., 2017b). However, environmental policies within many advanced and developing economics are increasingly being developed at the regional level to account for variations in economic and environmental needs (Fredriksson and Millimet, 2002; Prager and Freese, 2009; Prasad and Munch, 2012; Zhang and Wen, 2008). Therefore, there is a growing need for the use of EEIO models at subnational levels to support analyses that can inform regional environmental decision making. For regional EEIO analysis, a state or a province might be the most proper spatial resolution, because much of the official authority with regard to industrial and economic activity, as well as environmental oversight, resides at this level for most countries. States or provinces have an interest in encouraging economic growth, protecting public environmental health within their borders, and understanding the unique regional nature of activities within their jurisdiction.

Ideally, state-based or province-based EEIO models would be similar to the environmentally-extended multiregional input-output (EE-MRIO) models at the global scale (Tukker et al., 2009), with each state or province differentiated and connected through interregional commodity flows (Yang, 2016). Such a detailed model fully captures regional economic and environmental situations and interregional dependences. It can be used to study the national implications of regional economic activities or policies, for example, the life cycle environmental impacts of products consumed in one or any region and how the impacts are distributed in other regions (Isard, 1951). It can also be used to study the regional implications of national activities or policies, for example, the life cycle environmental impacts of an average product consumed in different regions and how each region may be affected (Yang and Heijungs, 2017). Developing such a multiregional EEIO model within a country, however, presents numerous challenges. Most states or provinces do not produce input-output tables. And trade may not be well tracked between states or provinces, in ways it is across international borders, because commodities flow freely within a country.

In addition, for policy makers of one state or province with limited resources and whose primary interest is their own jurisdiction, a detailed multiregional model may be impractical and beyond the scope. In this case, a more practical solution may be a simplified 2-region EEIO model, with one region being the state/province of interest and the other being the rest of the country. The simplified 2-region model can be relatively easily derived from a national EEIO model. Data only need to be collected on 1) the economic and environmental aspects of the region of interest and 2) trade between the region and the rest of the country, as opposed to collecting such data for all regions in the detailed multiregional EEIO model.

The question is, how accurate is the simplified 2-region model as opposed to the detailed multiregional EEIO model? Theoretically, the loss of spatial resolution could compromise the accuracy of the results. For example, the state/province of interest may purchase its products primarily from one or several states/provinces. Under these circumstances, the aggregation of all other states/provinces into one region may lead to over- or under-estimates of its supply chain impacts. On the other hand, the extent to which the accuracy may be compromised depends on a range of factors including how different the states/provinces are and how commodities flow between them.

We address these questions in this paper. Taking the United States (US) as a case study, we present a test of the accuracy of 2-region (state-rest of the country) EEIO models by comparing their results with that of a detailed multiregional EEIO model that differentiates all states and the District of Columbia. Our goal is to improve our understanding of the level of spatial resolution necessary for accurately modeling the life cycle environmental consequences of production or consumption activities within a state/province. Our study contributes to the regional LCA literature by exploring the important question of spatial scale or spatial aggregation in life cycle inventory (LCI) analysis. That is, how the numbers of regions and their sizes and industry profiles drive differences in LCI results. This subject remains largely unexplored (Yang, 2016; Yang and Heijungs, 2017), with the exception of Su and Ang (2010) using Chinese EEIO models. Studies of regionalized LCA have focused on life cycle impact assessment (LCIA), such as developing regionalized characterization factors (Hellweg and i Canals, 2014; Potting and Hauschild, 2006). Studies covering LCI analysis often regionalized the foreground processes only (Mutel et al., 2011; O’Keeffe et al., 2016; Tessum et al., 2012; Xue et al., 2015; Yang et al., 2012), or failed to account for linkages between regions (Yang, 2016).

2. Methods and data

We first develop a 51-region EEIO model based by regionalizing national input-output (IO) accounts. We then aggregate it into 51 unique 2-region (state-rest of the country) models. We compare the results from the 2-region models against that from the 51-region model, and calculate relative errors as an indication of how accurate the 2-region models are as a proxy for the 51-region model. The relative error is defined as:

$$RE = \frac{|m_1 - m_2|}{m_1} \quad (1)$$

where m_1 symbolizes life cycle inventory results from the 51-region model and m_2 results from the 2-region models. We explore reasons for variation of relative errors for different environmental flows. Details on computational structure and data compilation and processing are as follows.

2.1. Computational structure

The computational structure of the regionalized EEIO models is based on the use and make (UV) framework (Miller and Blair, 2009):

$$\mathbf{U} = \begin{bmatrix} \mathbf{U}^{1,1} & \mathbf{U}^{1,2} & \dots & \mathbf{U}^{1,n} \\ \mathbf{U}^{2,1} & \mathbf{U}^{2,2} & \dots & \mathbf{U}^{2,n} \\ \dots & \dots & \dots & \dots \\ \mathbf{U}^{n,1} & \mathbf{U}^{n,2} & \dots & \mathbf{U}^{n,n} \end{bmatrix} \quad (2)$$

$$\mathbf{g} = [\mathbf{g}^1 \quad \mathbf{g}^2 \quad \dots \quad \mathbf{g}^n] \quad (3)$$

$$\mathbf{V} = \begin{bmatrix} \mathbf{V}^{1,1} & \mathbf{V}^{1,2} & \dots & \mathbf{V}^{1,n} \\ \mathbf{V}^{2,1} & \mathbf{V}^{2,2} & \dots & \mathbf{V}^{2,n} \\ \dots & \dots & \dots & \dots \\ \mathbf{V}^{n,1} & \mathbf{V}^{n,2} & \dots & \mathbf{V}^{n,n} \end{bmatrix} \quad (4)$$

$$\mathbf{q} = [\mathbf{q}^1 \quad \mathbf{q}^2 \quad \dots \quad \mathbf{q}^n] \quad (5)$$

where $\mathbf{U}$ is the use table, reflecting commodities used by industries to produce their output. On-diagonal and off-diagonal blocks in $\mathbf{U}$ indicate intra- and inter-regional commodity flows. For example, $\mathbf{U}^{2,2}$ and $\mathbf{U}^{1,2}$ indicate commodities consumed in region 2 that are produced in region 2 and region 1, respectively. $\mathbf{V}$ is the make table, reflecting commodities produced by industries. Submatrices in $\mathbf{V}$ indicate contributions to commodities in a region. For example, $\mathbf{V}^{1,1}$ and $\mathbf{V}^{2,1}$ indicate total commodities available in region 1 that are from regions 1 and 2. $\mathbf{g}$ is a vector of total industry output, and $\mathbf{g}^i$ ($i=1, \dots, n$) indicates total output produced by industries in region i . And $\mathbf{q}$ is a vector of total commodity, and $\mathbf{q}^i$ ($i=1, \dots, n$) indicates total commodities available in region i .

In addition, total environmental emissions and resource use are expressed by (Suh et al., 2010):

$$\mathbf{E} = [\mathbf{E}^1 \quad \mathbf{E}^2 \quad \dots \quad \mathbf{E}^n] \quad (6)$$

where $\mathbf{E}^i$ ($i=1, \dots, n$) indicates direct emissions and resource use (per dollar) by industries in region i . To calculate the life cycle (cradle-to-gate) emissions and resources per dollar worth of a commodity (k) produced in region i :

$$\mathbf{m} = \mathbf{B}(\mathbf{I} - \mathbf{A})^{-1} \mathbf{f} \quad (7)$$

$$\mathbf{A} = \mathbf{U} \hat{\mathbf{g}}^{-1} \mathbf{V} \hat{\mathbf{q}}^{-1} \quad (8)$$

$$\mathbf{B} = \mathbf{E} \hat{\mathbf{g}}^{-1} \mathbf{V} \hat{\mathbf{q}}^{-1} \quad (9)$$

$$\mathbf{f} = \begin{bmatrix} 0 \\ \dots \\ 0 \\ (1)_k^i \\ 0 \\ \dots \\ 0 \end{bmatrix} \quad (10)$$

where the symbol “ $\hat{}$ ” indicates diagonalization; $\mathbf{V} \hat{\mathbf{q}}^{-1}$ is the market share matrix, and $\mathbf{A}$ is the direct requirement matrix (commodity-by-commodity); $\mathbf{E} \hat{\mathbf{g}}^{-1}$ and $\mathbf{B}$ symbolize direct emission and resource use intensities of industries and of commodities; $\mathbf{I}$ is the identity matrix, and $\mathbf{f}$ is the final demand vector, with all entries being 0 except for the commodity and region being studied.

In the 51-region EEIO model, n equals 51. And because the aggregate national IO accounts consist of 15 industries and 17 commodities, the dimension of $\mathbf{U}$, $\mathbf{V}$, $\mathbf{g}$, and $\mathbf{q}$ is 867 by 750, 750 by 867, 1 by 750, and 1 by 867 respectively. For $\mathbf{E}$, 10 air pollutants together with water withdrawal are covered in our analysis (section 2.2), so its 11 by 750. Then, the 51-region EEIO model can be aggregated into 512-region EEIO models (n equals 2), where the region of interest is singled out and all other regions aggregated to the rest of the country.

2.2. Data sources

The national IO accounts used as basis of the regional models are the 15-industry, 17-commodity make and use tables for year 2012, compiled by the Bureau of Economic Analysis (BEA, 2015). Note that imports are embedded in industry transactions in the national IO accounts. Therefore, life cycle inventory results estimated by the regional models include emissions and resource use occurring both domestically and internationally, assuming imports produced by the same technologies as used domestically. The effect of the role of the different import mixes of foreign goods in different states is not explored here, but other authors have explored consequence of different import modeling assumptions (e.g., Su and Ang 2013).

Inter- and intra-regional commodity flow ratios are derived from the 2012 commodity flow survey by the Bureau of Transportation Statistics and the U.S. Census Bureau (CFS, 2015). The survey covers primarily agricultural, manufacturing, and mining goods, and for services we assume they are provided locally, including electricity, commercial services, and governmental services. Note that the survey does not differentiate commodity flows intended

for use by industries, or by consumers, or commodity flows that represent intermediate shipments that might not be used within the region.

State industry output ratios are derived primarily from the industry revenue data in the 2012 economic census (USCB, 2016). The revenue data are generally available at the aggregate 2-digit NAICS (North American Industry Classification System) level. When such data are not available, other statistics such as payroll or number of establishments in the census are used as proxy. In addition, agricultural output ratios are derived from the 2012 annual cash receipts by commodity reported in the Farm Income and Wealth Statistics (USDA, 2016). Regional data on industry value-added provided by the BEA (2015) are used to approximate state-level government services.

Environmental satellite tables for each region cover a number of major air pollutants and water withdrawal. The air pollutants consist of $PM_{2.5}$, PM_{10} , ammonia, carbon monoxide, lead, sulfur dioxide, nitrogen dioxide, mercury, benzene, and cyanide. Data are from the 2011 national emissions inventory (NEI) compiled by the U.S. Environmental Protection Agency, which is publicly available (NEI, 2014). Following the general method of developing environmental satellite tables, as described in, e.g., Suh (2005) and Yang et al. (2017b), we map the emissions data from NEI to the industries in the state-level IO accounts described above. The water withdrawal data have been estimated from information compiled by the US Department of Agriculture (USDA) and US Geological Survey (USGS); see Yang et al. (2017b) for details. Table 1 summarizes the mean, median, standard deviation, and coefficient of variation of emission and resource intensity of industries in different states of the continental US. All the data compiled are available in a published dataset (Yang and Ingwersen, 2017).

2.3. Regionalization of the national make and use tables

The regionalization of the EEIO model presented here builds on Yang and Heijings (2016). Key elements used to regionalize the national IO accounts are 1) state industry output ratios and 2) intra- and inter-regional commodity flow ratios. Below is our step-wise approach to regionalization.

1. Calculate the direct requirement coefficients ($U\hat{G}^{-1}$ in equation 8) of the national use table, and apply them to each region;
2. Calculate state industry output (g in equation 3) using national industry output and regional output ratios derived from a mix of census data, farm receipts, and industry value added (section 2.2);
3. Estimate the 51-region Use table (U in equation 2) using 1), 2), and intra- and inter-regional commodity flow ratios from the commodity flow survey (section 2.2);
4. Calculate the market share matrix ($V\hat{q}^{-1}$ in Eq. 8 & 9) using the national make table, and apply them to each region;
5. Estimate total commodities available in different regions (q in equation 5) using intra- and inter-regional commodity flow ratios and regional industry output;

6. Estimate the 51-region Make table (V in equation 4) using 4) and 5), and intra- and inter-regional commodity flow ratios;

3. Results

As figure 1 reflects, the inventory results for the 2-region EEIO models roughly approximate that for the 51-region master model, with the median relative error smaller than 15% for all modeled emissions and resources except mercury (25%) and cyanide (23%). For water, $PM_{2.5}$, PM_{10} , sulfur dioxide, and benzene, the relative error is smaller than 20% for over 80% of the estimates. For ammonia, carbon monoxide, lead, nitrogen oxides, where the relative error is smaller than 20% for over 90% of the estimates. Results show that many of the high relative errors, as in water and benzene, are associated with agricultural and mining commodities, due to high regional variation in environmental intensities and active interstate trade.

Figure 2 explores the extent to which the accuracy of the 2-region EEIO models can be explained by the variability of environmental flows, measured by coefficient of variation (CV). CV is defined as the standard deviation of all direct intensity estimates for an environmental flow divided by the mean. Theoretically, the more variable an environmental flow is across regions, the greater the need for spatial differentiation. For example, if an industry is spatially homogenous and generates the same amounts of emissions wherever it is located, how it is spatially delineated would not affect the life cycle emissions of its purchasers. By contrast, if an industry is highly variable, its spatial delineation would play a role in its purchasers' life cycle emissions: the more refined the spatial resolution, the more accurate the purchaser's life cycle emissions. In other words, if an environmental flow is less spatially variable, we would find the 2-region models to be more accurate in approximating the 51-region model. Figure 2 confirms the theory, that is, a negative correlation between the variability of environmental flows and the accuracy of the 2-region models.

4. Discussion and Conclusions

In this study, we have conducted a test of the accuracy of simplified 2-region state EEIO models when compared to a detailed 51-region EEIO model in the context of the US. The results indicate the 2-region models seem to provide estimates for most of the regions, commodities, and environmental flows, with the relative error generally less than 20% of the detailed model. In accordance with theory, results show that variations in the accuracy of the 2-region models for the different environmental flows can be largely attributed to the spatial variability of the flow data.

The implications of our study are as follows. For researchers, industrial experts, and policy makers whose primary interest is their own state, the simplified 2-region EEIO model, with the state of interest being one region and the rest of the country the other region, may provide an adequate basis for preliminary or screening-level analysis, at least in the US. If the goal and scope of a study is about the total life cycle results, rather than about how the emissions or resource use are distributed across regions, the simplified 2-region model would be a suitable start. If more spatial resolution is required, the two-region model can be

further disaggregated into 3 regions, 4 regions, and so forth to improve model accuracy for parts of the life cycle as needed. Such spatial refinement is recommended if it is identified that most of the commodities used within an industry of the state of interest originate in only a few other states (regions) and those states differ greatly from other regions in terms of technology and environmental factors.

There are several limitations of our test that could be overcome in future studies, thus offering additional insights on the question of how the choice of spatial scale may affect LCI results. First, our test used highly-aggregated IO accounts. This aggregation may have smoothed out the effects of the variability observed for the compiled environmental flows and reduced the relative error calculated for them. However, the issue of aggregation depends on the goal and scope of a study. If a study investigates national or industry-level research questions, aggregate IO models may suffice (Su and Ang, 2010). On the other hand, if it investigates micro- or process-level research questions, aggregate IO models are likely to fall short (Yang et al., 2017a). The second potential limitation is related to our use of the national IO accounts to model each region in the absence of regional IO accounts. But this limitation should only have a small to moderate effect on the results on the following grounds. For some sectors such as manufacturing, production technologies may not vary significantly across regions, and for others that do such as agriculture, their variation is already captured in the environmental data. Note that for agriculture, the use of uniform production technologies would affect estimates of indirect emissions through intermediate inputs, but direct on-site emissions from input use usually dominate the life cycle emissions (Xue et al., 2015; Yang and Suh, 2015).

The third limitation is our assumption that, in the absence of trade data for services, all services are supplied from within the state. This assumption may suffice for some industries such as restaurant and dry cleaning, but may fall short for others such as finance. The effects of this assumption on the results, however, may be limited because the inventory for services is primarily attributed to their use of manufactured goods and agricultural commodities. If these goods are adequately modeled, estimates of the life cycle inventories of services should be reasonably accurate. The fourth limitation is the lack of accounting for trade in goods and services outside of national borders. If states are importing at different rates from different countries of origin, the interstate trade modeled here would not capture the emissions associated with any of these differences in import profiles between states. Given these potential limitations, the findings presented here should be cautiously interpreted as providing a measure of the accuracy of the simplified 2-region EEIO model. This paper intends to shed light on the relative accuracy of two forms of state-based US MRIO models; acceptability of the accuracy of the 2- and 51- region EEIO model results should be determined by the users based on the intended model use.

Finally, the analysis presented here should not be interpreted as a measure of the accuracy of regionalized modeling in LCA in general. We only evaluated the level of spatial resolution in the life cycle inventory (total emissions and resources used). If life cycle impact assessment (LCIA) methods are used to compute indicators of potential environmental impact, then regionalization in an LCA model could also include impact assessment models that incorporate regional variations in environmental conditions. For example, emitting a

kilogram of particulate matter in the middle of a densely-populated region can have more potentially detrimental human health effects than emitting that same kilogram in a remote, sparsely-populated region. Similarly, extracting one million liters of water from a depleted aquifer in an arid, water-scarce region can be more impactful than extracting that same quantity from a surface body in a water-rich region. Some studies have evaluated the effects of spatial resolution in LCIA methods for the US (e.g. (Henderson et al., 2017; Mutel et al., 2011)); nevertheless, such studies of impact assessment within the context of spatially-differentiated EEIO models of the US are needed.

References

- BEA. Input-Output Accounts Data [WWW Document]. 2015. URL http://www.bea.gov/industry/io_annual.htm
- CFS. [accessed 3.1.16] 2012 Commodity Flow Survey Public Use Microdata File [WWW Document]. 2015. URL <https://www.census.gov/econ/cfs/pums.html>
- Fredriksson PG, Millimet DL. Strategic Interaction and the Determination of Environmental Policy across U.S. States. *J. Urban Econ.* 2002; 51:101–122. DOI: 10.1006/juec.2001.2239
- Heijungs, R., Suh, S. The computational structure of life cycle assessment. Kluwer Academic Pub; Dordrecht: 2002.
- Hellweg S, i Canals LM. Emerging approaches, challenges and opportunities in life cycle assessment. *Science*. 2014; 344:1109–1113. [PubMed: 24904154]
- Henderson AD, Asselin-Balençon AC, Heller M, Lessard L, Vionnet S, Jolliet O. Spatial Variability and Uncertainty of Water Use Impacts from U.S. Feed and Milk Production. *Environ. Sci. Technol.* 2017; 51:2382–2391. DOI: 10.1021/acs.est.6b04713 [PubMed: 28068477]
- Isard W. Interregional and regional input-output analysis: a model of a space-economy. *Rev. Econ. Stat.* 1951:318–328.
- ISO. ISO 14040: Environmental management-Life cycle assessment-Principles and framework. International Organization for Standardization; Geneva, Switzerland: 2006.
- Kerkhof AC, Nonhebel S, Moll HC. Relating the environmental impact of consumption to household expenditures: An input–output analysis. *Ecol. Econ.* 2009; 68:1160–1170.
- Lenzen M. Primary energy and greenhouse gases embodied in Australian final consumption: an input–output analysis. *Energy Policy*. 1998; 26:495–506.
- McGregor PG, Swales JK, Turner K. The CO₂ “trade balance” between Scotland and the rest of the UK: Performing a multi-region environmental input–output analysis with limited data. *Ecol. Econ.* 2008; 66:662–673.
- Miller, R., Blair, P. Input-output analysis: foundations and extensions. Cambridge University Press; Cambridge, England: 2009.
- Mutel CL, Pfister S, Hellweg S. GIS-Based Regionalized Life Cycle Assessment: How Big Is Small Enough? Methodology and Case Study of Electricity Generation. *Environ. Sci. Technol.* 2011; 46:1096–1103. [PubMed: 22206467]
- Nansai, K. Environmental Input-Output Database Building in Japan. In: Suh, S., editor. *Handbook of Input-Output Economics in Industrial Ecology, Eco-Efficiency in Industry and Science*. Springer; Netherlands: 2009. p. 653–688.
- NEI. [accessed 11.14.16] National Emissions Inventory (NEI) [WWW Document]. 2014. URL <https://www.epa.gov/air-emissions-inventories/national-emissions-inventory-nei>
- O’Keeffe S, Wochele-Marx S, Thrän D. RELCA: aREgional Life Cycle inventory for Assessing bioenergy systems within a region. *Energy Sustain. Soc.* 2016; 6:12.doi: 10.1186/s13705-016-0078-8
- Potting J, Hauschild M. Spatial differentiation in life cycle impact assessment: a decade of method development to increase the environmental realism of LCIA. *Int. J. Life Cycle Assess.* 2006; 11:11–13.

- Prager K, Freese J. Stakeholder involvement in agri-environmental policy making – Learning from a local- and a state-level approach in Germany. *J. Environ. Manage.* 2009; 90:1154–1167. DOI: 10.1016/j.jenvman.2008.05.005 [PubMed: 18632199]
- Prasad M, Munch S. State-level renewable electricity policies and reductions in carbon emissions. *Energy Policy.* 2012; 45:237–242.
- Su B, Ang BW. Input–output analysis of CO₂ emissions embodied in trade: the effects of spatial aggregation. *Ecol. Econ.* 2010; 70:10–18.
- Suh S. Developing a sectoral environmental database for input–output analysis: the comprehensive environmental data archive of the US. *Econ. Syst. Res.* 2005; 17:449–469.
- Suh S, Weidema B, Schmidt JH, Heijungs R. Generalized Make and Use Framework for Allocation in Life Cycle Assessment. *J. Ind. Ecol.* 2010; 14:335–353.
- Tessum CW, Marshall JD, Hill JD. A spatially and temporally explicit life cycle inventory of air pollutants from gasoline and ethanol in the United States. *Environ. Sci. Technol.* 2012; doi: 10.1021/es3010514
- Tukker A, Poliakov E, Heijungs R, Hawkins T, Neuwahl F, Rueda-Cantuche JM, Giljum S, Moll S, Oosterhaven J, Bouwmeester M. Towards a global multi-regional environmentally extended input–output database. *Ecol. Econ, Methodological Advancements in the Footprint Analysis.* 2009; 68:1928–1937. DOI: 10.1016/j.ecolecon.2008.11.010
- USCB. [accessed 5.22.16] 2012 Economic Census [WWW Document]. 2016. URL <http://www.census.gov/econ/census/>
- USDA. [accessed 11.25.16] USDA ERS - Farm Income and Wealth Statistics [WWW Document]. 2016. URL <https://www.ers.usda.gov/data-products/farm-income-and-wealth-statistics/>
- Weber, C., Matthews, D., Venkatesh, A., Costello, CH., Scott, M. The 2002 US benchmark version of the economic input-output life cycle assessment (EIO-LCA) model. Green Design Institute, Carnegie Mellon University; Pittsburgh, PA: 2009.
- Wiedmann T. A review of recent multi-region input–output models used for consumption-based emission and resource accounting. *Ecol. Econ.* 2009; 69:211–222.
- Xue X, Hawkins TR, Ingwersen WW, Smith RL. Demonstrating an approach for including pesticide use in life-cycle assessment: Estimating human and ecosystem toxicity of pesticide use in Midwest corn farming. *Int. J. Life Cycle Assess.* 2015; 20:1117–1126. DOI: 10.1007/s11367-015-0902-y
- Yang Y. Toward a more accurate regionalized life cycle inventory. *J. Clean. Prod.* 2016; 112:308–315.
- Yang Y, Bae J, Kim J, Suh S. Replacing gasoline with corn ethanol results in significant environmental problem-shifting. *Environ. Sci. Technol.* 2012; 46:3671–3678. DOI: 10.1021/es203641p [PubMed: 22390573]
- Yang and Ingwersen. Simplified two and fifty-one region state-based EEIO model comparison [Data set]. US Environmental Protection Agency. 2017; doi: 10.23719/1390147
- Yang Y, Heijungs R. A generalized computational structure for regional life-cycle assessment. *Int. J. Life Cycle Assess.* 2017; 22:213–221. DOI: 10.1007/s11367-016-1155-0
- Yang Y, Heijungs R, Brandão M. Hybrid life cycle assessment (LCA) does not necessarily yield more accurate results than process-based LCA. *J. Clean. Prod.* 2017a; 150:237–242. DOI: 10.1016/j.jclepro.2017.03.006
- Yang Y, Ingwersen WW, Hawkins TR, Srocka M, Meyer DE. USEEIO: A new and transparent United States environmentally-extended input-output model. *J. Clean. Prod.* 2017b; 158:308–318. DOI: 10.1016/j.jclepro.2017.04.150
- Yang Y, Suh S. Changes in environmental impacts of major crops in the US. *Environ. Res. Lett.* 2015; 10:94016.
- Yang Y, Suh S. Environmental Impacts of Products in China. *Environ. Sci. Technol.* 2011; 45:4102–4109. [PubMed: 21469663]
- Zhang K, Wen Z. Review and challenges of policies of environmental protection and sustainable development in China. *J. Environ. Manage.* 2008; 88:1249–1261. DOI: 10.1016/j.jenvman.2007.06.019 [PubMed: 17767999]

Figure 1.

A Tukey boxplot of the relative errors in using aggregated 2-region EEIO models as proxies of the detailed 51-region EEIO model. Results represent the continental US, with Hawaii, Alaska, and District of Columbia excluded. There are in total 816 data points—48 regions multiplied by 17 commodities—for each environmental flow. Whiskers represent the 1.5 interquartile range (IQR) from the lower or higher quartile, and any data points outside the range are identified as outliers.

Figure 2.

Correlation between variability of environmental flows across regions and industries, as measured by the coefficient of variation (CV), and the accuracy of the 2-region EEIO models, as measured by mean relative errors. The larger the CV, the more variable an environmental flow is across regions or industries. Relative errors are calculated by equation (1) as an indication of how accurate the 2-region models are.

Table 1

Descriptive statistics on emission and resource use intensity (kg/\$) of industries across regions.

	Mean	Median	Standard deviation	Coefficient of variation
Water	9.07E+01	1.11E+00	4.34E+02	479%
Ammonia	5.17E-04	7.96E-08	1.98E-03	382%
Carbon Monoxide	3.49E-03	1.30E-05	1.09E-02	312%
Lead	6.07E-08	2.62E-11	2.04E-07	335%
Nitrogen Oxides	1.49E-03	1.65E-05	4.62E-03	309%
PM2.5	3.91E-04	3.55E-06	1.29E-03	330%
Sulfur Dioxide	1.15E-03	1.42E-06	7.25E-03	631%
Benzene	1.21E-05	4.48E-07	3.80E-05	315%
Cyanide	3.15E-07	0.00E+00	1.91E-06	608%
PM10	2.23E-03	6.93E-06	1.07E-02	480%
Mercury	9.14E-09	3.87E-12	5.98E-08	654%