

A generalized computational structure for regional life-cycle assessment

Yi Yang^{1,2} · Reinout Heijungs^{3,4}

Received: 1 February 2016 / Accepted: 14 June 2016 / Published online: 28 June 2016
© Springer-Verlag Berlin Heidelberg 2016

Abstract

Purpose Regional life-cycle assessment (LCA) is gaining an increasing attention among LCA scholars and practitioners. Here, we present a generalized computational structure for regional LCA, discuss in-depth the major challenges facing the field, and point to a direction in which we believe regional LCA should be headed.

Methods Using an example, we first demonstrate that when there is regional heterogeneity (be it due to environmental conditions or technologies), average data would be inadequate for estimating the life-cycle impacts of a product produced in a specific region or even that of an average product produced in many regions. And when there is such regional heterogeneity, an understanding of how regions are connected through commodity flows is important to the accuracy of regional LCA estimates. Then, we present a generalized computational structure for regional LCA that takes into account interregional commodity flows, can evaluate various cases of regional differentiation, and can account for multiple impact categories simultaneously. In so doing, we show what kinds of data are required for this generalized framework of regional LCA.

Results and discussion We discuss the major challenges facing regional LCA in terms of data requirements and computational complexity, and their implications for the choice of an optimal regional scale (i.e., the number of regions delineated within the geographic boundary studied).

Conclusions We strongly recommend scholars from LCI and LCIA to work together and choose a spatial scale that not only adequately captures environmental characteristics but also allows inventory data to be reasonably compiled or estimated.

Keywords Commodity flow · Heterogeneity · Regional life-cycle assessment · Spatial differentiation · Spatial scale

1 Introduction

Spatial differentiation in life-cycle assessment (LCA) arose two decades ago in response to recognitions of the problems with applying site-generic life-cycle impact assessment (LCIA) models to evaluating non-global impact categories such as human health respiratory illness (Potting and Hauschild 2006). From a life-cycle perspective, a product is linked to many processes located in different places with likely different characteristics such as population density. Site-generic LCIA neglects regional conditions and derives generic, global characterization factors (CFs) for chemicals released anywhere within the geographic boundary studied (e.g., globe or a country). When there is substantial environmental heterogeneity, however, using site-generic CFs may be highly problematic. It would be clearly erroneous, for example, to suggest that releasing 1 kg of benzene in Hong Kong would cause the same damage to human health as releasing it in rural Tibet. Therefore, developing spatially differentiated CFs can improve the relevance and accuracy of LCA results. Over the

Responsible editor: Rolf Frischknecht

✉ Yi Yang
yyandgin@gmail.com

¹ CSRA Inc., 3170 Fairview Park Drive, Falls Church, VA 22042, USA

² Bren School of Environmental Science and Management, University of California, Santa Barbara, USA

³ Department of Econometrics and Operations Research, Vrije Universiteit Amsterdam, Amsterdam, The Netherlands

⁴ Department of Industrial Ecology, Institute of Environmental Sciences, Leiden University, Leiden, The Netherlands

past two decades, many spatialized LCIA models have been developed, and there seems to be a trend of increasingly refined spatial scale toward grid-cell-based CFs (Heijungs 2012; Yang 2016).

However, LCIA is not the only phase where spatial differentiation is needed. It may also improve the relevance and accuracy of life-cycle inventory (LCI) analysis. Similar to generic CFs in LCIA, average data are often used in LCI analysis. But average data do not provide an adequate basis for either global or regional impact categories when there is substantial spatial heterogeneity in technology. Such heterogeneity is evident across nations (Lenzen and Wachsman 2004; Mutel et al. 2013), but may also exist at the subnational level. This is especially true for agricultural systems (Yang et al. 2012; Yang and Suh 2015). Technology variability across regions within a country may be caused by many factors, including differences in environmental conditions as in the case of agriculture (Miller et al. 2006; Chiu et al. 2009) and in the level of development as in the case of China (Du et al. 2012). Recently, there has been a large body of LCA research with spatial differentiation at the inventory level, mostly focused on biofuel and agricultural products (Tessum et al. 2012; Yang et al. 2012; Geyer et al. 2013; Yang and Suh 2015).

Basically there are two ways of spatial differentiation: on the basis of geographical regions (e.g., states, provinces, watershed, grid cells; see e.g., (Bare et al. 2003; Wegener Sleswijk and Heijungs 2010; Núñez et al. 2015) or on the basis of characteristics (e.g., urban vs. rural, clay soil vs. peat soil; see e.g., (Frischknecht and Rebitzer 2005; Rosenbaum et al. 2008)). In this paper, we concentrate on the first form of spatial differentiation, where a geographically defined boundary demarcates two regions. We will use the term “regional LCA” to refer to this form of spatial differentiation, and we emphasize at the outset that it can be carried out at the inventory phase or impact phase or both. Our first goal is to demonstrate that there is a commonality between regional differentiation due to environmental conditions and that due to technologies. That is, when there is regional heterogeneity—be it due to technologies or the environment, average data would be inadequate for estimating the life-cycle impacts of a product produced in a specific region or even that of an average product produced in many regions. And when there is such regional heterogeneity, an understanding of how regions are connected through commodity flows is important to the accuracy of regional LCA estimates.

Our second goal is to present a generalized computational structure for conducting regional LCA. It is generalized in the sense that it takes into account interregional commodity flows, can evaluate various cases of regional differentiation, and can account for multiple impact categories simultaneously, as opposed to previous methods that can accommodate only a

single emission or a single impact category (see, e.g., (Yang 2016)) or evaluate only a single process instead of the whole life cycle (see., (Mutel et al. 2011)). Also, we present data requirements for this generalized framework of regional LCA as compared to data requirements for the standard, unregionalized LCA that the mainstream LCA community has been working on so far. Next, we discuss in detail implications of this framework and the major challenges facing regional LCA. Last, we point to a direction in which we believe regional LCA should be headed.

2 A generalized framework for regional life-cycle assessment

In this section, we use a simple, fictitious two-process example of wheat-based beer production to show that there are various cases of regional differentiation where regionalization is needed and can improve the accuracy and relevance of results. The regional differentiation may be due to heterogeneity of technology or environmental factors. We seek to demonstrate that the different cases of regionalization can all be modeled by a single framework that takes into account interregional commodity flows. We are guided by two types of research questions, namely, the life-cycle environmental impacts of a product produced in a specific region or in all regions within the geographic boundary studied. The relevance of the first type of question is reflected in, e.g., policies in a specific region that have affected other regions, and that of the second type of question is reflected in, e.g., national policies that have affected all regions in the country. The two types of questions may be further categorized into global impacts such as climate change or regional impacts such as human health respiratory effect.

2.1 Unregionalized LCA

First, let us briefly review the standard procedure for LCA, introducing a simple site-generic example of beer production. For this we will resort to matrix-based notation (cf. (Heijungs and Suh 2002)), as it is easily generalizable to regional LCA. Suppose country X produces wheat beer, among many other products, and the beer's life cycle consists of beer and wheat production, both of which are produced in country X . To estimate the life-cycle environmental impacts of beer, data were collected on how beer and wheat were made. We learned that to make 1 liter of beer, 5 kg of wheat was used, and as a result 0.4 kg of CO₂, 0.8 kg of NO_x, and 1.7 g PM₁₀ were released; and to produce 1 kg of wheat, no inputs were used but 1.5 kg of CO₂, 2 kg of NO_x, and 1 g of PM₁₀ were released. There is also a set of mid-point characterization factors (CFs) that has been developed for the country showing that the climate change impact of 1 kg CO₂ is 1 kg CO_{2e}, the acidification

impact of 1 kg NO_x in the country is 1.2 kg H⁺e., and the human health respiratory impact of PM₁₀ and NO_x in the country is 0.6 and 0.1 g PM_{2.5}e, respectively. A standard matrix-based formulation (Heijungs and Suh 2002) of the example is as follows.

$$\mathbf{A} = \begin{bmatrix} 1 & -5 \\ 0 & 1 \end{bmatrix} \tag{1}$$

$$\mathbf{B} = \begin{bmatrix} 1.5 & .4 \\ 2 & .8 \\ 1 & 1.7 \end{bmatrix} \tag{2}$$

$$\mathbf{f} = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \tag{3}$$

$$\mathbf{Q} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1.2 & 0 \\ 0 & .1 & .6 \end{bmatrix} \tag{4}$$

A is the technology matrix, in which columns represent wheat and beer production and rows represent wheat (in kg) and beer (in liter), respectively. **B** is the environmental matrix, in which rows represent CO₂ (in kg), NO_x (in kg) and PM₁₀ (in g) emissions, respectively. **f** represents the functional unit of a study, which is 1 liter of beer in our example. And **Q** is CF matrix, with rows representing climate change, acidification, and respiratory impact categories (in kg CO₂e, kg H⁺e, and g PM_{2.5}e respectively), and columns representing CO₂, NO_x, and PM₁₀ emissions (in the units mentioned before). Now we can calculate the life-cycle emissions of beer as follows.

$$\mathbf{m} = \mathbf{BA}^{-1}\mathbf{f} = \begin{bmatrix} 7.9 \\ 10.8 \\ 6.7 \end{bmatrix} \tag{5}$$

where **m** shows that the life-cycle CO₂, NO_x, and PM₁₀ emissions of 1 liter of beer are estimated at 7.9 kg, 10.8 kg, and 6.7 g, respectively. Then we can continue to calculate the life-cycle impacts of beer with the following equation:

$$\mathbf{h} = \mathbf{Qm} = \begin{bmatrix} 7.9 \\ 13.0 \\ 5.1 \end{bmatrix} \tag{6}$$

where **h** shows that the life-cycle climate change, acidification, and human health respiratory impacts of 1 liter of beer in country *X* are estimated at 7.9 kg CO₂e, 13.0 kg H⁺e, and 5.1 g PM_{2.5}e. In summary, the data and LCA analysis above reflect a typical country-based LCA study, where process data are sectoral averages, there is little locational information on the processes involved, and CFs are country-based without considering regional environmental conditions. As we shall show in next sections, such generic data and LCA analysis falls short if there is technological or/and environmental heterogeneity within the geographic boundary studied.

2.2 Distinguishing technological and environmental heterogeneity

Below, we use Fig. 1 to illustrate the different cases of heterogeneity due to either technology or the environment or both. Again, our goals are to (1) demonstrate the inadequacy of average data in the face of regional heterogeneity and show the importance of understanding how regions are connected through commodity flows, and (2) present a generalized computational structure for regional LCA that addresses the various cases of regional heterogeneity.

2.3 Environmental heterogeneity only

Suppose there is environmental heterogeneity in country *X* while technologies are homogenous. For example, population density varies significantly, and based on population density,

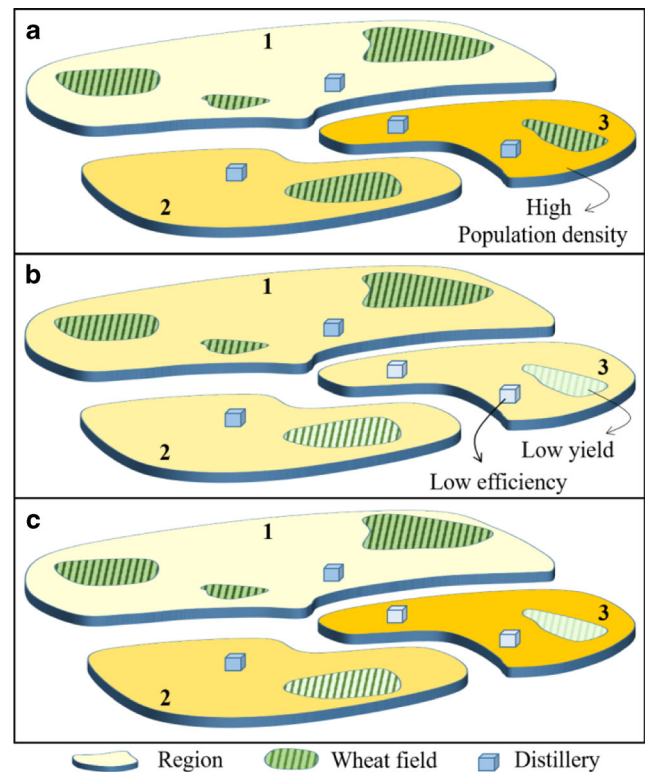


Fig. 1 A simple, fictitious two-process, three-region example of wheat beer production. Note that in this example besides beer, wheat is also used to make other products such as bread. **a** Technological homogeneity and environmental heterogeneity, where wheat and beer production technologies are identical in all regions but population density differs, with region 3 having the highest density as reflected by the *dark yellow color*. **b**: Technological heterogeneity and homogenous environment, where population density is identical across regions but region 3 has lower wheat yield and lower beer production efficiency than the other two regions as reflected by the *lighter colors*. **c** Technological and environmental heterogeneity, where both regions differ in population density, wheat yield, and beer production efficiency

the country can be divided roughly into 3 regions, with region 3 having the highest population density followed by region 2 and then region 1 (Fig. 1a). Suppose we are further given location information on all wheat fields and beer distilleries as mapped out in Fig. 1.

In this case, the original inventory we compiled and CF model built (Eqs. 1–5) become inadequate for estimating the life-cycle impact of beer produced in any single region and even for the average beer in the country. There are two dimensions of inadequacy that need to be stressed. First, based on the original CF model, the impact of beer production would be identical anywhere it was produced. But because of the variation in population density, beer made in region 3 is likely to have generated higher life-cycle impacts than beer brewed in region 2. This impact-related inadequacy has been well recognized and facilitated regional LCA development in the past two decades (Potting et al. 1998; Potting and Hauschild 2006).

The other dimension of inadequacy is related to LCI. The original inventory compiled is a portrayal of *generic* wheat used by all industries not the wheat used *specifically* by distilleries, and because population density varies, an understanding of where wheat was purchased from is important because it may change the results. For example, the life-cycle impact of beer production in this country would be lower if distillers purchased wheat from region 1 than if they purchased wheat from region 3. This inventory-related inadequacy was recently pointed out (Yang 2016). The two inadequacies discussed, however, are specific to NO_x and PM_{10} only, because they contribute to regional environmental impacts and are subject to regional conditions. For CO_2 , location does not matter: wherever it was produced, it would generate the same CO_2 emissions (due to the assumption of technological homogeneity) with the same climate impact.

To address the inadequacies, we need to collect data on (1) wheat flows between regions, (2) regional output volume of wheat, and (3) region-specific CFs. Suppose we found out that 80 % of wheat used in distilleries in region 3 was from within the region and 20 % from region 1, 60 % of wheat used in distilleries in region 2 was from within the region and 40 % from region 1, and 100 % of wheat used in distilleries in region 1 was from within the region. Also, we learned that beer output was 20, 30, and 50 liters in regions 1–3. Further, the acidification impact to the country due to NO_x emissions from regions 1–3 is 1.1, 1.2, and 1.6 kg H^+e per kg NO_x , respectively; the respiratory impact to the country due to NO_x from regions 1–3 is 0.09, 0.12, and 0.15 g $\text{PM}_{2.5\text{e}}$ per kg NO_x respectively; and that due to PM_{10} from regions 1–3 is 0.5, 0.7, and 0.9 g $\text{PM}_{2.5\text{e}}$ per g PM_{10} .

With such information, we can proceed to modify the site-generic framework by creating submatrices that reflect region-

specific processes, products, environmental emissions, and CFs, shown as follows,

$$\mathbf{A}_r = \begin{bmatrix} \mathbf{A}^{11} & \mathbf{A}^{12} & \mathbf{A}^{13} \\ \mathbf{A}^{21} & \mathbf{A}^{22} & \mathbf{A}^{23} \\ \mathbf{A}^{31} & \mathbf{A}^{32} & \mathbf{A}^{33} \end{bmatrix} = \begin{bmatrix} 1 & -5 & 0 & -2 & 0 & -1 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 1 & -3 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 1 & -4 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \quad (7)$$

$$\mathbf{B}_r = \begin{bmatrix} \mathbf{B}^{11} & \mathbf{B}^{12} & \mathbf{B}^{13} \\ \mathbf{B}^{21} & \mathbf{B}^{22} & \mathbf{B}^{23} \\ \mathbf{B}^{31} & \mathbf{B}^{32} & \mathbf{B}^{33} \end{bmatrix} = \begin{bmatrix} 1.5 & .4 & 0 & 0 & 0 & 0 \\ 2 & .8 & 0 & 0 & 0 & 0 \\ 1 & 1.7 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 1.5 & .4 & 0 & 0 \\ 0 & 0 & 2 & .8 & 0 & 0 \\ 0 & 0 & 1 & 1.7 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 1.5 & .4 \\ 0 & 0 & 0 & 0 & 2 & .8 \\ 0 & 0 & 0 & 0 & 1 & 1.7 \end{bmatrix} \quad (8)$$

$$\mathbf{f}_r = \begin{bmatrix} \mathbf{f}^1 \\ \mathbf{f}^2 \\ \mathbf{f}^3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0.2 \\ \hline 0 \\ 0.3 \\ \hline 0 \\ 0.5 \end{bmatrix} \quad (9)$$

$$\mathbf{Q}_r = \begin{bmatrix} \mathbf{Q}^1 & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{Q}^2 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{Q}^3 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1.1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & .09 & .5 & 0 & 0 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1.2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & .12 & .7 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1.6 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & .15 & .9 \end{bmatrix} \quad (10)$$

where subscript r serves to distinguish the new matrixes with regional differentiation from the original ones and dashed lines in Eqs. 7–10 serve to partition and differentiate regions. In Eq. 7, on-diagonal (e.g., A^{22}) and off-diagonal blocks (e.g., A^{23}) partitioned by the dashed lines indicate intraregional and interregional commodity flows, respectively. For example, column 4 in Eq. 7 indicates that to produce 1 liter of beer in region 2, 5 kg of wheat was used, of which 3 kg was supplied from within the region and 2 kg from region 1. Now we can calculate the regional life-cycle emissions and impacts of beer in this country as follows,

$$m_r = B_r A_r^{-1} f_r = \begin{bmatrix} 3.2 \\ 4.4 \\ \hline 2.4 \\ 1.5 \\ \hline 2.0 \\ \hline 1.4 \\ \hline 3.2 \\ 4.4 \\ 2.9 \end{bmatrix} \tag{11}$$

$$h_r = Q_r m_r = \begin{bmatrix} 3.2 \\ 4.8 \\ \hline 1.6 \\ 1.5 \\ \hline 2.4 \\ \hline 1.2 \\ \hline 3.2 \\ 7.0 \\ 3.2 \end{bmatrix} \tag{12}$$

where m_r indicates that the amounts of, e.g., NO_x from regions 1–3 are estimated at 4.4, 2.0, and 4.4 kg, and h_r indicates that regions 1–3 contribute 4.8, 2.4, and 7.0 kg H^+e to acidification impact in this country attributed to 1

liter of beer production. We can sum across emissions and impacts from different regions as follows,

$$m = I_r m_r = \begin{bmatrix} 7.9 \\ 10.8 \\ 6.7 \end{bmatrix} \tag{13}$$

$$h = I_r h_r = \begin{bmatrix} 7.9 \\ 14.3 \\ 6.1 \end{bmatrix} \tag{14}$$

$$I_r = [I \quad I \quad I] \tag{15}$$

where the submatrices in I_r are identity matrices, the dimension of which depends on the number of emissions (as in Eq. 13) or impact categories (as in Eq. 14). The results show that the total amounts of emissions attributed to beer production remain unchanged as compared to the unregionalized example, because of the assumption of technology homogeneity. But because of environmental heterogeneity, the life-cycle impacts of beer now are different from that calculated based on average data (compare Eqs. 14 and 6). We can also adjust the final demand vector and evaluate the life-cycle impacts of beer produced in a single region, for example, in region 3.

$$f_r = \begin{bmatrix} 0 \\ 0 \\ \hline 0 \\ 0 \\ \hline 0 \\ 0 \\ 1 \end{bmatrix} \tag{16}$$

$$m = \begin{bmatrix} 7.9 \\ 10.8 \\ 6.7 \end{bmatrix} \tag{17}$$

$$h = \begin{bmatrix} 7.9 \\ 16.3 \\ 7.1 \end{bmatrix} \tag{18}$$

2.4 Technological heterogeneity only

Suppose there is only technological heterogeneity in country X while the environment is homogenous. Let us assume, for example, that wheat yield varies greatly with region 3 having the lowest yield and hence the highest rates of emissions (Fig. 1b). Also, beer production in region 3 is less efficient than in the other 2 regions, e.g., 6 gram of wheat used to produce 1 liter of beer in region 3 (Fig. 1b). Details on the technological and environmental aspects are provided below (cf., Eqs. 7 and 8).

$$A_r = \begin{bmatrix} A^{11} & A^{12} & A^{13} \\ A^{21} & A^{22} & A^{23} \\ A^{31} & A^{32} & A^{33} \end{bmatrix} = \begin{bmatrix} 1 & -5 & 0 & -2 & 0 & -1.2 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & -3 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & -4.8 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \quad (19)$$

$$B_r = \begin{bmatrix} 1.2 & .4 & 0 & 0 & 0 & 0 \\ 1.6 & .8 & 0 & 0 & 0 & 0 \\ .8 & 1.7 & 0 & 0 & 0 & 0 \\ 0 & 0 & 2 & .4 & 0 & 0 \\ 0 & 0 & 3 & .8 & 0 & 0 \\ 0 & 0 & 1.5 & 1.7 & 0 & 0 \\ 0 & 0 & 0 & 0 & 3 & .5 \\ 0 & 0 & 0 & 0 & 4 & 1 \\ 0 & 0 & 0 & 0 & 2 & 2 \end{bmatrix} \quad (20)$$

In this case, the original inventory (Eqs. 1–3 and 5) we compiled becomes inadequate also, for precisely the same reason discussed above, namely, a lack of understanding of where wheat was supplied from. This is because now where wheat was produced matters and can change the results. Likewise, to address this inadequacy, we need to compile data on wheat flows between regions.

Using assumptions regarding regional output percentages as reflected in Eq. 9 in section 2.1., we can

calculate the life-cycle emissions of 1 liter of beer produced in country X using Eqs. 11 and 13.

$$m_r = \begin{bmatrix} 2.7 \\ 3.7 \\ \frac{2.1}{1.9} \\ 2.9 \\ \frac{1.9}{7.5} \\ 10.1 \\ 5.8 \end{bmatrix} \quad (21)$$

$$m = \begin{bmatrix} 12.1 \\ 16.7 \\ 9.8 \end{bmatrix} \quad (22)$$

Note that the results are much higher than that in previous two examples (cf., Eqs. 13 and 5). This is mainly due to the fact that in our example region 3 is least efficient in producing wheat and beer, while it has the largest market share of beer. Of course, regionalized emissions could also have been lower with other data. The same applies to the life-cycle impacts (Eqs. 23–24), even given the assumption of environmental homogeneity (Eq. 25).

$$h_r = Q_r m_r = \begin{bmatrix} 2.7 \\ 4.4 \\ \frac{1.6}{1.9} \\ 3.5 \\ \frac{1.4}{7.5} \\ 12.1 \\ 4.5 \end{bmatrix} \quad (23)$$

$$h = \begin{bmatrix} 12.1 \\ 20.1 \\ 7.5 \end{bmatrix} \quad (24)$$

$$\mathbf{Q}_r = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1.2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & .1 & .6 & 0 & 0 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1.2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & .1 & .6 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1.2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & .1 & .6 \end{bmatrix} \tag{25}$$

Likewise, we can calculate the life-cycle emissions and impacts of 1 liter of beer produced in region 3 only (Eqs. 26–27). Note that the results are even higher than that of 1 average liter of beer produced (Eqs. 22 and 24). Again, this is because of the low wheat yield and low beer production efficiency in region 3 and because distilleries purchased wheat primarily from within the region.

$$\mathbf{m} = \begin{bmatrix} 16.3 \\ 22.1 \\ 12.6 \end{bmatrix} \tag{26}$$

$$\mathbf{h} = \begin{bmatrix} 16.3 \\ 26.5 \\ 9.7 \end{bmatrix} \tag{27}$$

2.5 Both technological and environmental heterogeneity

Suppose there is heterogeneity of both technologies and environmental conditions in country *X*. That is, population density, wheat yield, and beer production efficiency all vary across regions. This third case is the most general one, and in fact the most likely in real-world case studies. In this case, as discussed above, both the original built inventory and impact model are inadequate. To address the inadequacies, we need (1) a region-specific inventory, (2) regional output volumes of beer, (3) interregional commodity flows, and (4) region-specific CFs.

Using data from the previous sections (Eqs. 9, 10, 19, and 20), we can calculate the life-cycle emissions and impacts of 1 liter of beer produced in country *X* (Eqs. 28–29, and compare with Eqs. 13, 14, 22, and 24).

$$\mathbf{m} = \begin{bmatrix} 12.1 \\ 16.7 \\ 9.8 \end{bmatrix} \tag{28}$$

$$\mathbf{h} = \begin{bmatrix} 12.1 \\ 23.7 \\ 9.8 \end{bmatrix} \tag{29}$$

2.6 A generalized framework for regional LCA

As mentioned, the third case of both technological and environmental heterogeneity is the most likely in the real world, and the former two can be considered special cases. Therefore, a generalized framework we propose for regional LCA consists of the following elements:

- Region-specific unit process data
- Regional process output volumes
- Interregional commodity flows
- Region-specific CFs

The generalized computational structure is as follows,

$$\mathbf{h}_r = \mathbf{Q}_r \mathbf{B}_r \mathbf{A}_r^{-1} \mathbf{f}_r \tag{30}$$

$$\mathbf{Q}_r = \begin{bmatrix} \mathbf{Q}^1 & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \mathbf{Q}^2 & \dots & \mathbf{0} \\ \dots & \dots & \dots & \dots \\ \mathbf{0} & \mathbf{0} & \dots & \mathbf{Q}^j \end{bmatrix} \tag{31}$$

$$\mathbf{B}_r = \begin{bmatrix} \mathbf{B}^{11} & \mathbf{B}^{12} & \dots & \mathbf{B}^{1j} \\ \mathbf{B}^{22} & \mathbf{B}^{22} & \dots & \mathbf{B}^{2j} \\ \vdots & \vdots & \dots & \vdots \\ \mathbf{B}^{j1} & \mathbf{B}^{j2} & \dots & \mathbf{B}^{jj} \end{bmatrix} \tag{32}$$

$$\mathbf{A}_r = \begin{bmatrix} \mathbf{A}^{11} & \mathbf{A}^{12} & \dots & \mathbf{A}^{1j} \\ \mathbf{A}^{22} & \mathbf{A}^{22} & \dots & \mathbf{A}^{2j} \\ \dots & \dots & \dots & \dots \\ \mathbf{A}^{j1} & \mathbf{A}^{j2} & \dots & \mathbf{A}^{jj} \end{bmatrix} \tag{33}$$

$$\mathbf{f}_r = \begin{bmatrix} \mathbf{f}^1 \\ \mathbf{f}^2 \\ \dots \\ \mathbf{f}^j \end{bmatrix} \tag{34}$$

where subscript *j* denotes *j* regions delineated within the geographic boundary studied. Total impacts across regions can be calculated as follows. Notice that each of the elements of \mathbf{h}_r , \mathbf{Q}_r , \mathbf{B}_r , \mathbf{A}_r , and \mathbf{f}_r is a vector or matrix as well. Thus, even though \mathbf{B}_r might appear square, \mathbf{B}^{11} , etc., may be non-square, from which \mathbf{B}_r inherits a non-square structure. It is also worth stressing that \mathbf{B}_r reflects emissions of origin such that off-diagonal submatrices are all 0 s (see, e.g., Eq. 8), and \mathbf{Q}_r reflects impacts of emissions from one region to the entire country (or all regions delineated). This means that the characterization matrix (\mathbf{Q}_r) accounts for the

fate and transport of emissions across regions, in addition to effects (Rosenbaum et al. 2007). The total can be calculated by,

$$\mathbf{h} = \mathbf{I}_r \mathbf{h}_r \quad (35)$$

$$\mathbf{I}_r = [\mathbf{I} \quad \mathbf{I} \quad \dots \quad \mathbf{I}] \quad (36)$$

3 Discussion and conclusions

Thus far, we have demonstrated how LCA has been done traditionally, how regionalization can improve the accuracy of results when there is environmental and/or technological heterogeneity, and how different circumstances of regionalization can all be evaluated by a generalized framework that differentiates processes, products, and CFs by region, and takes into account interregional commodity flows. Part of the reason we present such a generalized framework is to discuss its implications for the choice of an optimal regional scale for regional LCA.

The major challenge facing regional LCA is a mismatch between LCIA and LCI (Mutel et al. 2011; Hellweg and i Canals 2014), that is, the two having different regional scales. Such a mismatch, however, may take different forms. For example, LCI and LCIA have roughly the same number of regions, but region shapes may be drastically different (Fig. 2a). The regional scale for LCI may be based on geopolitical boundaries (e.g., state or province), whereas that for LCIA may be derived from ecotypes such as population density and watershed (Mutel et al. 2011; Hellweg and i Canals 2014). Or LCIA has a much more refined regional scale than LCI (Fig. 2c). An example

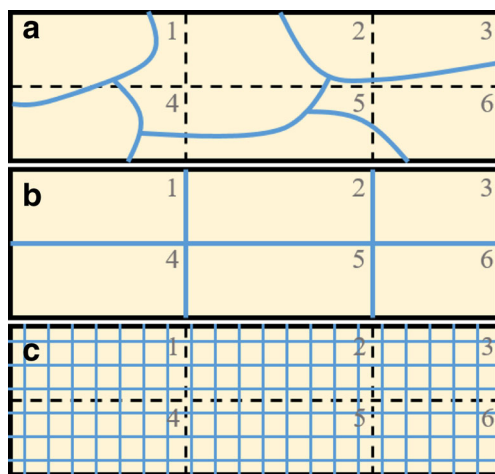


Fig. 2 Common regional scales used for LCI and LCIA, illustrated by an example of a country consisting of six regions as delineated by the dashed lines (i.e., geopolitical division). Inventory data exist at both the national and regional level. **a** Mismatch of regional scale between LCI and LCIA, where *blue lines* reflect the regional scale for LCIA (based on, e.g., population density or watershed). **b** The same regional scale for LCI and LCIA. **c** Grid-cell-based regional scale for LCIA

of the sort is grid-cell-based LCIA, which seems to become increasingly popular (Heijungs 2012).

Before we elaborate on the challenge posed by this mismatch, let us examine the easiest case where LCIA has the same regional scale as LCI does, based on geopolitical division such as state, as inventory data often exist at this level (Fig. 2b) (Wegener Sleeswijk and Heijungs 2010). There are, for example, data on agricultural use by state collected by the US Department of Agriculture (USDA) and also state-level CFs for certain impact categories developed by the Environmental Protection Agency (EPA) (Bare et al. 2003). In this case, region-specific inventory can be easily connected with region-specific CFs. However, this is not the end of the calculation. Because none of the regions exist in vacuum and instead they may be highly connected through trade as shown in previous sections, we need data on interregional commodity flows to fully understand a product's supply-chain impacts across regions. But as such interregional flows may not be readily available at a fine resolution (e.g., county-level), neither may inventory data, our choice of regional scale in this case is limited. We may, for example, do a state-level regional LCA after collecting inter-state commodity flows data, but we may not be able to do a county-level regional LCA because neither county-level inventory data nor inter-county commodity flows data are available even if county-specific CFs are available or can be easily developed.

In the case of a mismatch between LCI and LCIA, therefore, it is even more challenging to conduct a complete regional LCA. In Fig. 2a, for example, what would be the direct environmental impacts of product Y produced in region 4 if the industry producing Y are scattered across region 4? The difficulty multiplies as we begin to think about product Y's supply-chain emissions and impacts, which likely span all regions. Even if we can map out all the industries, regroup them, and approximate a new set of region-specific inventories aligned with the regional scale of LCIA (i.e., blue lines in Fig. 2a), we are still faced with the question of interregional commodity flows. They may not be easily adapted to the regional scale of LCIA (i.e., blue lines in Fig. 2a), even if such data are available on the regional scale of LCI (i.e., dashed lines in Fig. 2a). The challenge becomes extremely thorny when LCIA takes a grid-cell-based regional scale (Fig. 2c). It is of great difficulty to derive grid-cell-based inventory data for all industries, and even if that could be done, there are no data on inter-grid-cell commodity flows. The essence of life cycle breaks down on this level of regional resolution even though grid-cell-based CFs may be more accurate than state or county-based CFs from an impact analysis perspective. It is telling that most grid-cell-based LCIA proposals do not contain an illustrative example that includes even a part of a real life cycle (cf. Heijungs 2012).

The idea of LCA is a powerful one, which provides us with a holistic view of the problems at hand and allows us to

identify “hotspots” that are worthwhile focusing on. In fact, however, what LCA does is split every industry, with their emissions and environmental impacts, and attribute the proportions to different functional units, or services in accordance to input/output relationships. In regional LCA, therefore, the task then becomes to puzzle out the proportion of every industry in every region that is attributable or related to the functional unit being studied. Region-specific inventory and interregional commodity flows are the key to estimating such regional proportions, but the more regions defined, the harder it becomes to collect data for these two, to store and make available these data in supersized matrices, and to do the required calculations. Matrix inversion is typically a $n^2 \log(n)$ process, which implies that a 10 times bigger **A** matrix requires a >200 times longer computation time (Heijungs et al. 2015). For future work on regional LCA, we urge scholars from LCI and LCIA working together and choosing a regional scale that not only adequately captures environmental characteristics but also allows inventory data to be reasonably compiled or estimated.

References

- Bare J, Norris G, Pennington D, McKone T (2003) The tool for the reduction and assessment of chemical and other environmental impacts. *J Ind Ecol* 6:49–78
- Chiu Y, Walseth B, Suh S (2009) Water embodied in bioethanol in the United States. *Environ Sci Technol* 43:2688–2692
- Du L, Wei C, Cai S (2012) Economic development and carbon dioxide emissions in China: provincial panel data analysis. *China Econ Rev* 23:371–384
- Frischknecht R, Rebitzer G (2005) The ecoinvent database system: a comprehensive web-based LCA database. *J Clean Prod* 13:1337–1343
- Geyer R, Stoms D, Kallalos J (2013) Spatially-explicit life cycle assessment of sun-to-wheels transportation pathways in the US. *Environ Sci Technol* 47:1170–1176
- Heijungs R (2012) Spatial differentiation, GIS-based regionalization, hyperregionalization and the boundaries of LCA. In: Ioppolo G (ed) *Environment and energy*. Milano, Italy, pp 165–176
- Heijungs R, Suh S (2002) *The computational structure of life cycle assessment*. Kluwer Academic Pub, Dordrecht
- Heijungs R, de Koning A, Wegener Sleeswijk A (2015) Sustainability analysis and systems of linear equations in the era of data abundance. *J Environ Account Manag* 3:109–122
- Hellweg S, i Canals LM (2014) Emerging approaches, challenges and opportunities in life cycle assessment. *Science* 344:1109–1113
- Lenzen M, Wachsmann U (2004) Wind turbines in Brazil and Germany: an example of geographical variability in life-cycle assessment. *Appl Energy* 77:119–130
- Miller S, Landis A, Theis T (2006) Use of Monte Carlo analysis to characterize nitrogen fluxes in agroecosystems. *Environ Sci Technol* 40:2324–2332
- Mutel CL, Pfister S, Hellweg S (2011) GIS-based regionalized life cycle assessment: how big is small enough? Methodology and case study of electricity generation. *Environ Sci Technol* 46:1096–1103
- Mutel CL, de Baan L, Hellweg S (2013) Two-step sensitivity testing of parametrized and regionalized life cycle assessments: methodology and case study. *Environ Sci Technol* 47:5660–5667
- Núñez M, Pfister S, Vargas M, Antón A (2015) Spatial and temporal specific characterisation factors for water use impact assessment in Spain. *Int J Life Cycle Assess* 20:128–138
- Potting J, Hauschild M (2006) Spatial differentiation in life cycle impact assessment: a decade of method development to increase the environmental realism of LCIA. *Int J Life Cycle Assess* 11:11–13
- Potting J, Schöpp W, Blok K, Hauschild M (1998) Site-dependent life-cycle impact assessment of acidification. *J Ind Ecol* 2:63–87
- Rosenbaum RK, Margni M, Jolliet O (2007) A flexible matrix algebra framework for the multimedia multipathway modeling of emission to impacts. *Environ Int* 33:624–634
- Rosenbaum RK, Bachmann TM, Gold LS et al (2008) USEtox—the UNEP-SETAC toxicity model: recommended characterisation factors for human toxicity and freshwater ecotoxicity in life cycle impact assessment. *Int J Life Cycle Assess* 13:532–546
- Tessum CW, Marshall JD, Hill JD (2012) A spatially and temporally explicit life cycle inventory of air pollutants from gasoline and ethanol in the United States. *Environ Sci Technol* 46(20):11408–11417
- Wegener Sleeswijk A, Heijungs R (2010) GLOBOX: a spatially differentiated global fate, intake and effect model for toxicity assessment in LCA. *Sci Total Environ* 408:2817–2832
- Yang Y (2016) Toward a more accurate regionalized life cycle inventory. *J Clean Prod* 112:308–315
- Yang Y, Suh S (2015) Land cover change from cotton to corn in the USA relieves freshwater ecotoxicity impact but may aggravate other regional environmental impacts. *Int J Life Cycle Assess* 20:196–203
- Yang Y, Bae J, Kim J, Suh S (2012) Replacing gasoline with corn ethanol results in significant environmental problem-shifting. *Environ Sci Technol* 46:3671–3678