A Review of Methods Used to Aggregate Distinct Impact Categories in LCA Studies

Mark Smurthwaite*

Liben Jiang

Karl S. Williams

School of Engineering

University of Central Lancashire

Preston, PR1 2HE

*Corresponding authors: Mark Smurthwaite MSmurthwaite3@uclan.ac.uk

Abstract

An in-depth literature review was carried out, highlighting previous life cycle assessments (LCA) studies on the environmental impact of food packaging solutions.

Purpose: The aim was to analyse the different indicators used and the relationships thereof, the various methods weighing distinct damage categories and the implicit trade-offs between the various environmental burdens.

Method: The results of 40 studies were analysed. The frequency of comparisons between different impact categories was counted. The frequency of correlation between impact categories was measured in 270 comparisons as well as the trade-off levels in instances where changes in impact were not correlated. Methods by which authors combined distinct categories were reviewed.

Results and discussion: It was found that the most frequently used indicators were correlated in >93% of comparisons. Authors used subjective or often undefined methods of combining and comparing distinct categories.

Conclusion: Methods used to combine distinct impact or damage categories can be affected by subjective value judgment and can generate arbitrary comparisons. Suggestions are made to improve such methods.

Keywords: Sustainability; food packaging; Life Cycle Assessment; environmental impacts; trade-offs; literature review.

1. Introduction

1.1 LCA Methodology

In the field of Life Cycle Assessment (LCA), results are given in many ways, using different methods with which to find them (British Standards Organisation, 2018). This is a concern for decision-makers who must interpret results from multiple assessments of distinct impact categories (e.g. global warming potential, primary energy demand) which are complex and often poorly represented, with key steps to support decisions being neglected (Björklund, 2002; Sala and Andreasson, 2018). This article is concerned with two potential solutions: (1) Measurement of a single impact category and (2) Application of a universal impact unit aggregating all impact categories available.

The single impact category solution has received some support (Frischknecht et al., 2015), though some debate continues as to which would be the most suitable category, and how it is measured, with different methods returning significantly different results (Arvidsson and Svanstrom, 2015). While calculating a product's impact in terms of carbon equivalent emissions and other impact categories provides valuable information for decision-makers, it has two drawbacks. Firstly, it can be expensive, time-consuming and resource-intensive potentially causing LCA not to be conducted where it might be useful (Puig et al., 2013). Secondly, it is dependent on knowledge or assumptions about more variables including the private and public vehicle fleet and the energy mix in electricity supply, which can vary by a factor of 160 between two countries (Fruergaard et al., 2009). Its dependence on such inputs makes results vulnerable to obsolescence as these inputs change over time or in different places (Björklund, 2002). An approach that measures the cumulative energy demand (CED) of a system – an impact category used within LCA which considers the energy extracted from nature used for energy and production – offers solutions to both concerns (Arvidsson and Svanstrom, 2015). CED accounting for the total energy harvested for a product ensures consistency and independence from an unknown and changing energy mix allowing for wider and more long-term applicability (Frischknecht et al., 2015). Furthermore, CED has been shown to be a reliable predictor of environmental impact (Huijbregts et al., 2006; 2010; Puig et al., 2013).

The first aim of this article is to investigate the appropriateness of measuring only energy demand as a means of comparing environmental impact.

The aggregated impact category solution is more difficult to implement, not only because of the increased scope of measuring multiple impacts, but also the distinctness of such categories which must be aggregated into a common value using subjective value judgements. Such judgments used in LCA studies affect the results, and should be guided by the values of the decision-maker, rather than person conducting the LCA (Miettinen and Hämäläinen, 1997).

Previous methods have suggested weighting as a way of addressing this challenge (Goedkoop and Spriensma, 1999). As per ISO standards (British Standards Organisation, 2018), some methods extrapolate the inventory categories into midpoint, endpoint and aggregated impact scores in an attempt to simplify results for comparison. Converting the endpoints, or in some cases the midpoints, to a single score involves the controversial process of weighting damage categories; a process which is carried out inconsistently, allowing subjective reason to influence results. "The weighting process is not technical, scientific, or objective […] not necessarily based on natural science but often on political or ethical values", (Jensen et al., 1997, p.67). The distinct impact category results before weighting cannot be described as free from subjective interpretation (Goedkoop and Spriensma, 1999), but the choices to use weighting or valuation, and which method to use are both subject to ideology and personal intuition, further corrupting the data (Finnveden, 1997).

The second aim of this article is to investigate the utilisation of aggregated impact scores and other instances of subjective value judgments being used to compare distinct impact categories.

1.2 LCA of Food Packaging

To understand the effect of the LCA method on results, a literature review was carried out looking at food packaging LCA studies. Food packaging suggests itself as an interesting area for an investigation of this kind due to the various interconnected, and often politically charged, environmental concerns involved which would benefit from a consolidation of distinct goals.

There have been countless papers published investigating the environmental impacts of food packaging, with many focusing on the direct negative impacts (e.g. Geyer et al., 2017; Rahimi and Garcia, 2017) such as microplastic build-up in the environment (e.g. Wang et al., 2019; Nalbone et al., 2021), acidification of natural resources, depletion of fossil fuels, destruction of animal habitats and others focusing on the indirect benefits such as the reduction of foods loss and waste (FLW) through extension of perishable food shelf-life, protection during transport and efficient stock replenishment procedures (e.g. Wikström et al., 2014). While some of these papers present positive and negative outcomes in several environmental impact indicators, they rarely evaluate the relative importance of each impact or suggest an acceptable trade-off between them.

To my knowledge, there has not been an in-depth review of such studies which has analysed the selection of impact indicators, the relationship between the results, and compiled implicitly and explicitly approved trade-off levels expressed by the authors, as this article aims to do.

2. Method

2.1. Literature Search

The Scopus database was searched for journal articles containing specific words in their title, abstract and keywords. The Scopus search criteria is shown below:

(TITLE-ABS ("Food" OR "beverage" W/ "packaging" OR "bottle" AND "LCA" OR "life cycle assessment" OR "climate change" OR "global warming" OR "food waste" OR "plastic waste" OR "human toxicity" OR "acidification" OR "resource depletion" OR "eutrophication" OR "particulate matter" OR "environmental impact") DOCTYPE (ar) PUBYEAR > 1999 AND PUBYEAR < 2021 PUBSTAGE (final))

This search returned 284 articles to be reviewed. 244 were removed as they were either not concerned with assessing environmental impacts, did not contain a comparison, quantitative data or only measured a single impact such as CO2 emissions, leaving 40 for further analysis.

2.2. Analysis of Relationship Between Energy Demand and Other Impact Categories

The number of bilateral comparisons in each of the 41 articles was calculated from the number of products being compared in each (2 scenarios = 1 comparison; 3 scenarios = 3 comparisons; 4 scenarios = 6 comparisons etc.).

The impact categories measured in each comparison were recorded to determine the most frequently used. The frequency of each category being measured in combination with energy demand was counted to determine which relationship could be analysed with largest data set. The combinations with the most instances were analysed to determine how frequently they were correlated.

Where the 2 impact categories had an inverse relationship, the trade-off value between (the cost of one to save the other) was calculated by dividing the difference in energy demand across 2 products by the difference in the other category across the same products (trade-off = Δ_E / $\Delta_{category2}$). The range of trade-off values for each combination as well as any scenario recommendation given by the author(s), implicitly evaluating the acceptability of such a trade-off, was recorded.

2.3 Analysis of Methods of Aggregating Results

The articles which used a single aggregated category were reviewed to uncover the method by which different categories were converted into a single unit and whether the value judgments were explained sufficiently to allow interpretation of the results for decision makers.

It was noted where impact categories in their original units were given as these are essential for decision-makers using the data from the various studies. Where normalisation is stated to have been used, it is noted whether the method by which this has been done is explained. Finally, it is noted whether weighing of normalised impacts has been stated to have been used, and in such cases it is noted whether the method by which this has been done is sufficiently explained. While explanation of the objective impact values in their appropriate units is ideal for use by decision-makers, a thorough explanation of normalisation and weighing is a sub-optimal compromise, enabling some understanding of the data, while a single aggregated unit with no explanation of normalisation or weighing is of little use to decision-makers.

3. Results and Discussion

Of the 41 studies reviewed, 29 presented the results in the distinct categories only. The other 11 gave an aggregated score based on one of many methods used to consolidate distinct categories: Eco-indicator 99 (4), CML2001 (1), Impact 2002+ (4), and ReCiPe (2).

3.1 Selection frequency of each impact category

Table 1 shows the frequency of inclusion of each impact category used in more than one of the 29 articles reviewed. Several of the studies compare the environmental impact of more than two products, resulting in multiple direct product comparisons. The total number of comparisons in which each impact category is used is also given in the table. The studies which presented the energy demand impact category were reviewed to assess the category's applicability as an indicator of overall environmental impact. The number of comparisons which included energy demand and each other impact category is shown in the last column of the table.

Table 1: Various impact categories used in LCA studies

Impact category	**Abbreviation**	**Unit**	**Inclusions in studies**	**Total comparisons**	**Total comparisons which also included En**
Global warming potential, climate change, GHG emissions	GWP	CO_2 eq.	29	278	107
Acidification potential, acidification (terrestrial, aquatic)	AP	SO_2 eq., mol H^+ eq., species/year	20	144	88
Photochemical oxidant creation potential, photochemical ozone formation	POC	C_2H_4 eq., NMVOC eq., ethane-eq., NO_x eq.	15	171	75
Ozone depletion	ODP	R11 eq., CFC-11 eq., DALY	14	113	74
Eutrophication potential	EP	PO_4 eq., mol N eq.	12	115	81
Primary Energy Use, Cumulative energy demand, energy use (renewable, non-renewable), resource consumption	En	MJ	12	107	NA
Freshwater ecotoxicity	FWE	1,4-DCB eq., species/year, CTUe	8	95	3
Human toxicity potential	HTP	1,4-DB eq., C_2H_4 eq., DALY, DCB eq.	7	77	9
Particulate matter	PM	$PM_{2.5}$ eq., DALY, Disease incidence	7	21	7
Fossil fuels	Fos	oil eq., $	6	72	1
Terrestrial ecotoxicity	TE	1,4-DCB eq., species/year	6	28	3
Marine aquatic ecotoxicity	MaE	1,4-DCB eq., species/year	6	28	3
Water resource depletion, water consumption	WRD	m^3	6	14	9
Freshwater eutrophication	FEu	P eq., species/year	6	12	6
Abiotic Depletion Potential, Mineral, fossil & ren resource depletion	ADP	Sb eq.	4	25	3
Non-carcinogens	NCar	1,4-DCB eq., toluene eq.	4	21	9
Carcinogens	Car	1,4-DCB eq., benzene eq.	4	13	0
Mineral resource	Min	Cu eq.	4	12	9
Ionizing radiation	IR	kBq Co60 eq., DALY	4	5	0

Land use	LU	C deficit, m2a crop eq., m^2	4	4	0
Marine eutrophication	MEu	N eq	3	4	0
Human toxicity (soil)	HTs	m^3 soil, kg	2	13	3
Respiratory effects	Res	kg, $PM_{2.5}$ eq.	2	13	3
Ecotoxicity	ETX	2,4-D eq	2	11	0
Smog	Sm	NOx eq., O3-eq.	2	11	0
Organic respiratory compounds	OrC	%	2	10	9
Human health	HH	PM_{10}-eq., DALY	2	7	6
Metal depletion	Met	Fe eq., $	2	3	0

GWP is the only impact category which was measured in all the studies. While measurement of GWP is standard practice in LCA studies, inconsistent measurement of other categories makes it difficult to compare results of different studies, and raises questions about the validity of comparisons which do not include some important impact categories.

3.2 Relationship Between Energy Demand and Other Impact Categories

The difference in energy demand and the other categories in each comparison was reviewed to uncover the frequency of correlation, and the ratio of deltas in cases where the categories had an inverse relationship.

Energy demand v GWP

Across all articles review, energy demand and GWP were compared for 107 products or scenarios. They had a direct relationship in 101 (94.3%) of those comparisons. The six scenarios in which there was a conflict between the two impact categories are illustrated in Table 2. The different trade-off values (the energy cost required to reduce GWP or vice versa), the mean value and standard deviation (σ) are shown in the table.

Table 2: Trade-off (MJ/kg CO2) values in various studies

Comparison	**Δ_E Energy demand (MJ)**	**Δ_G GWP (kg CO2 eq.)**	**Trade-off Δ_E/Δ_G (MJ/kg CO2)**	**Trade-off acceptable by authors?**
Bohlmann, 2004	1036	3	**345.333(high)**	Not stated
Guiso et al., 2016_1	1.494	0.087	17.1724	Not stated
Guiso et al., 2016_2	1.526	0.049	31.1429	Not stated
Guiso et al., 2016_3	0.554	0.127	4.3622	Not stated
Saleh, 2016	2964.3	2113.58	**1.4025(low)**	Authors recommend lower energy demand. <1.4025
Zhang et al., 2015	0.284	0.005	56.8	Not stated
		Mean	76.0	
		σ	133.5	

Energy demand v Acidification Potential

Of the 88 comparisons between energy demand and acidification potential, 85 (96.6%) showed a direct relationship between the 2 categories. The different trade-off values (Δ_E/Δ_{AP}), the mean value and standard deviation (σ) are shown in Table 3.

Table 3: Trade-off (MJ/g SO2) values in various studies

Comparison	**Δ_E Energy demand (MJ)**	**Δ_{AP} AP (g SO2 eq.)**	**Trade-off Δ_E/Δ_{AP} (MJ/g SO2)**	**Trade-off acceptable by authors?**
Guiso et al., 2016	0.019	0.388	**0.049(low)**	Not stated
Vidal et al., 2007	4.5	1.58	**2.848(high)**	Authors recommend lower energy demand. <2.8481
Zhang et al., 2015	0.284	0.2	1.42	Not stated
		Mean	1.439	
		σ	1.14	

Energy demand v Eutrophication potential

Of the 81 comparisons between energy demand and eutrophication potential, 76 (93.8%) showed a direct relationship between the 2 categories. The different trade-off values (Δ_E/Δ_{EP}), the mean value and standard deviation (σ) are shown in table 4.

Table 4: Trade-off (MJ/g PO4) values in various studies

Comparison	**Δ_E Energy demand (MJ)**	**Δ_{EP} EP (g PO4 eq.)**	**Trade-off Δ_E/Δ_{EP} (MJ/g PO4)**	**Trade-off acceptable by authors?**
Vidal et al., 2007	4.5	0.391	11.509	Authors recommend lower energy demand. <11.509
Zhang et al., 2015	0.284	0.1	2.84	Not stated
Bertolini et al., 2016	0.7	0.0118	**59.322(high)**	Not stated
Guiso et al., 2016_1	0.051	1.494	**0.034(low)**	Not stated
Guiso et al., 2016_2	0.145	0.554	0.261	Not stated
		Mean	14.793	
		σ	25.325	

The three distinct impact categories most commonly measured in combination with energy demand (GWP, AP and EP) had a direct relationship with energy demand in more than 93% of comparisons. Where impacts had an inverse relationship, ratios between the impact categories varied widely across studies. Comparisons between products returned energy (MJ) to GWP (kg CO2 eq.) ratios between 1.4 and 345 with a high standard deviation (see table 2). This is because the studies are comparing different products using different LCA boundaries. It should be expected, but demonstrates how a stated acceptable trade-off value could be used by decision-makers with multiple areas of interest (e.g. policy makers, large diverse producers) to more effectively attain their environmental goals. Without using a standardised trade-off value as a guide, the combined choices of independently acting decision-makers could result in an increase in both damage categories overall. For example, if one decision-maker using the data from Bohlmann (2004) chose the reduced GWP product, and another decision-maker using the data from Saleh (2016) chose the least energy intensive product (as was recommended by the author) the result would be a net increase in both impact categories because of the disproportionate trade-off values in each case. A similar disparity between the trade-off ratio can be seen against the other categoies.

3.2 Application of the Aggregated Results Methods

12 gave articles used an aggregated score based on one of many methods used to consolidate distinct categories: Eco-indicator 99 (4), CML2001 (1), Impact 2002+ (4), and ReCiPe (2).

Eco-Indicator 99

The Eco-Indicator 99 method (Goedkoop and Spriensma, 1999, 2001) models a product's lifecycle impacts by calculating the midpoint impacts in various categories, converts them into three damage categories and weights them into a single score based on value judgements of individuals.

The Eco-Indicator 99 method of weighting different impact categories is based on the responses of several panel members, who were asked to weigh the relative importance of human life loss, extinction of plant species and depletion of resources, three categories which can be abstract and difficult to compare for individuals, to discover a collective ethic within a society (Goedkoop and Spriensma, 1999, 2001). Aside from the inherent difficulty of quantitatively valuing a human life compared to the extinction of a plant species, the authors of the system admit that the panel used were not representative of the society and that the responses varied so widely that the average value returned from the responses had limited meaning. According to the authors, the weight of the various impacts is subjective, does not have a single truth and should not be used for recommendations to the public.

Before weighting is even considered the method introduces subjectivity via the inclusion of "fundamental, or model uncertainties […] caused by unavoidable ethical and thus value based choices." (Goedkoop and Spriensma, 1999, p.15), including time-frame, value-based discounting, level of proof and whether to include damages which are more manageable. Each of these considerations "have significant effects on the result." (p.15). The method was used in four LCA studies in the survey (De Monte et al., 2005; Banar and Cokaygil, 2008; Bugnicourt et al. 2013; Cabot et al., 2019).

De Monte et al. (2005) did not explain how weighting was applied in their study of coffee packaging, but simply presented their results firstly, as inventory items in their individual units (kWh, MJ, g PET in production, g CO2 in certain processes etc.) and secondly, as combined normalized midpoint categories. Presenting the results in such a way does not allow a decision maker to apply their own valuation or even understand those of the author.

Banar and Cokaygil (2008) used the hierarchical perspective of the method and gave results in the three damage categories, but only presented the normalized values and did not explain how weighting was applied. In this study of juice packaging no weighting would make a difference to the overall comparison as one product performed better in all damage categories, hence there was no need for normalisation or aggregation of damage points.

Bugnicourt et al. (2013) used the normalised values with the egalitarian weighting method, one of three suggested in the Eco-Indicator methodology report (Goedkoop and Spriensma, 1999, 2001). The three weighting paradigms: Individualist, hierarchical and egalitarian were explained in the methodology report to respond to the inherent subjectivity of data interpretation, with each paradigm taking a unique approach to burden of proof, timescale, value judgement discounting and other subjective choices. As in Banar and Cokaygil (2008), one product performed more favourably in all damage categories so there was no need to apply the weighting.

Similarly, Cabot et al. (2019) only gave the LCA results as normalised damage points and did not explain how this was done or how weighting was applied. This study was not comparing multiple products, but simply assessing the comparative impacts of a single products' different components. The outcome, in this case, was affected by the choice of weighting. The PET component was determined to have the greatest impact based primarily on its impact on resources. Had human health been given greater weighting in the comparison, three other components (data cable, sensor and microcontroller) might have had greater overall impacts.

CML 2001

The CML method (Guinée et al. 2002) does not specifically recommend any weighting, although it suggests that "The criteria used to evaluate impact categories should be described extensively." (p.629), and that weighting should be avoided wherever possible. They also state that, as per ISO140 (British Standards Organisation, 2018), weighting factors, pre-normalisation data, value choices and their rationale shall be included to ensure that trade-offs are available to decision-makers.

One of the studies in the review used the CML method. Conte et al. (2015) presented their LCA results of various cheese packaging solutions simply as a single unit for each product and scenario. The individual impact categories' values were not given and the method of producing the single indicator was explained in 1 sentence: "The normalization and weighting of the results were performed by using the standard values in GaBi functionalities in order to determine an eco-indicator for each scenario." (Conte et al., 2015 p.13). Presenting the results in such a way is contrary to the cited method and the ISO standard on which it is based and diminishes the utility of the results for decision-makers.

Impact 2002+

Adapted from Eco-Indicator 99 and CML2002, the Impact2002+ method was developed with an additional midpoint indicator, climate change/global warming (Jolliett et al., 2003). The method's utility for comparing multiple products or processes is limited in that, unlike Eco-Indicator, it does not propose a weighting for the four damage categories. The authors suggest using one's own self-determined weighting factors if aggregation is needed. An updated user guide (Humbert et al., 2012) suggests using the method of De Schryver et al. (2009) to convert the effects of global warming into one or more of the other damage categories, to reduce the total number to be considered. While the weighting option in Eco-Indicator is far from optimal, having none or allowing subjective weighting allows personal biases to affect the LCA results. From the literature review, four articles (Siracusa et al., 2014; Ingrao et al., 2015; Ingrao et al., 2017 and Joachimiak-Lechman et al., 2018) used Impact 2002+ damage points to display their LCA results.

One stage of the Impact method is normalisation, in which the four damage categories are converted to damage points per emission. According to Jolliett et al. (2003, p.329), "The normalized factor is determined by the ratio of the impact per unit of emission divided by the total impact of all substances of the specific category for which characterization factors exist per person per year.". The normalization factor is a result of comparing the impact of a product or service to the background damage of the average person in the country or region of study. Applying this factor and combining the points without considering weighing means that damage categories are seen as less harmful in regions with higher background damage in that category. As a result, an LCA in a region with higher carbon emissions would consider a product with higher carbon emissions to be less harmful than a region with lower carbon emissions. Four studies in the literature review used the Impact LCA method.

Siracusa et al. (2014) used the Impact method and carried out normalization and weighing, but did not specify how each stage was done and did not present the raw data which would allow another comparison to be carried out using alternative methods of weighing.

Ingrao et al. (2015) gave their results as both midpoint and endpoint indicators in their impact assessment of the of the various inputs to produce a foam tray. Citing Siracusa et al. (2014), they claim that this allows different damage and impact categories to be easily compared to each other. No weighting method was explained, making the results difficult to interpret. However, it appears from the figures illustrating the difference in damage points, one product performed better in all categories, rendering the normalisation and weighting process redundant.

Ingrao et al. (2017) citing Lo Giudice et al. (2017) multiplied the normalised results 1 point (effectively applying equal weighting to the results). Stating the weighing method, while not sufficient for making a meaningful comparison, at least allows a decision-maker to apply their own valuations.

Joachimiak-Lechman et al. (2018) do not explain how the different damage categories are weighed, simply presenting the combined impacts as "points", but it is assumed that they are weighted equally after normalisation. The lack of explanation, as in Siracusa et al. (2014), restricts decision-makers from using the data and applying their own valuation.

Combining the normalised damage points and applying an equal weighting, as with each of the four studies, implies that the existing damage being done in all categories is of equal significance. This implication has an inherent value judgement and has an effect on the overall result. It also creates a paradox whereby any changes to the ambient environmental damage categories, through the implementation of a given solution, will affect the normalisation factors, which will in turn affect the result of the LCA and the preferability of that solution (i.e. as background damage in one category increases, solutions which would reduce such damage becomes less preferable; when damage in a certain category approaches negligible levels, solutions which would reduce such damage would be the most preferable.). This is obviously contrary to the intended result of impact assessments and sustainability goals.

Recipe

The Recipe 2017 method is the updated version of the method first developed by Goedkoop et al. (2009) (Huijbregts et al., 2016). Derived from Eco-Indicator 99, it converts midpoint impact categories to endpoint areas of protection, addressing areas of uncertainty and value judgments with the inclusion of three unique perspectives grounded in cultural theory. Unlike the other methods described it does not suggest a method for weighing the midpoint or endpoint categories into an aggregated points total. Two studies in the survey cited this method of measuring environmental impact.

Bertoluci et al. (2014) gave results of a comparison of different olive packing solutions in points only. Using the Hierarchist perspective and European Weighting, results are presented as a graph indicating different endpoint impacts excluding their specific values and a graph indicating the combined damage "Pts". Had the endpoint graph included exact values, it would be possible to apply different perspective and weightings. As the data is presented, it would only be useful for decision-makers in Europe who have decided to also utilise the hierarchist perspective.

Fresán et al. (2019) used the method and presented the midpoint results as points rather than distinct units (e.g. CO2 equivalent emissions, PM2.5 equivalent particulate matter). As with Cinelli et al. (2017), the authors did not explain the normalisation phase, weighting or which of the three perspectives is used. If it were stated, as must be assumed, that the geographically appropriate normalisation factors are used with equal weighting being given to each impact, it would be possible to revert the results to their distinct units for each category. While the necessity to do this brings into question the utility of normalisation and weighting in any case other than at the very point of decision, the exclusion of the conversion factors renders the results data useless.

Miscellaneous Valuation Methods

Aggregated indicators have been shown to involve subjectivity in the best examples, and in many cases cannot be used by decision-makers. For decision-makers, it is preferable to obtain results as separate endpoint damage categories, and more preferable still are midpoint impact categories. 29 of the studies in the review used this approach, but three included suggestions based on the results.

Vidal et al. (2007) state that conventional petroleum-derived plastic packaging has a more detrimental impact on the environment than a biodegradable, starch-derived packaging despite it performing better in 2 of the 4 categories measured. They determine this through normalisation without specifying a weighting. While the conclusion is not supported by objective data, reporting the individual impacts at least allows a decision-maker to apply their own valuation.

Bertolini et al. (2016) do not state a recommendation outright but begin their conclusion by stating that a multilayer milk carton has a lower environmental impact in most categories measured than a HDPE bottle. This may have no bearing on the comparative environmental impacts depending on the priorities of a decision-maker. The individual impact categories and their appropriate units are presented, allowing decision-makers to interpret the results for themselves.

Saleh (2016) ranked the environmental impacts of PET bottles, glass bottles and aluminium cans using a subjective weighting whereby the impact in each category was multiplied by a number from 1 to 10 based on the perceived importance of each category by a panel. There are several flaws in this method. Firstly, it disproportionately focuses on categories which are measured using higher numbers in their measurement. Respiratory effects, for example, are given in values between 0.1 and 0.7kg, whereas non-renewable energy is given in values between 300 and 12,300MJ. It is obvious that comparing these values without normalisation, as the author did, will mean that no realistic amount of respiratory effects would affect the outcome of the comparison. Secondly, multiplying these values by numbers from 1 to 10 implies that one category is 10 times more significant than one of the others. The result of this method meant that a glass bottle was stated to have a lower impact on the environment than an aluminium can despite having more than 3 times the level of human toxicity, 5 times the global warming potential, 7 times the terrestrial acidification potential and producing 9 times as much solid waste. Unlike some of the other studies, however, Saleh presented the midpoint data to allow a third party to make a decision, disregarding the arbitrary weighting applied.

4. Conclusions

4.1. Relationships between distinct impact categories

The high correlation between energy demand and the other categories (>93%) suggest that it may be useful for making interim time-critical decisions, but due to the potentially high trade-off values shown in this review, should not be relied upon as an indicator of environmental impact for long-term decisions, where the opportunity to carry out a more thorough LCA is available.

Often a decision-maker will not have the technical knowledge to appropriately aggregate distinct categories themselves, or appreciate the assumptions and uncertainties in pre-aggregated results. It would therefore be advisable to make available a standardised trade-off value between two impact categories as a guide, limiting the combined negative net effect of two independent decision-makers acting in consideration of their conflicted interest and biases. The large disparity between trade-off ratios in the studies reviewed in this paper highlight the potential jeopardy of making decisions without such guidance.

4.2 Weighting categories for direct comparison

Converting the midpoint categories to damage categories is a useful method of simplifying results although it does introduce subjectivity as it converts the different impacts into a few uniform units (human health, ecosystem quality, resources and climate change/global warming). Furthermore, De Schryver's et al. (2009) method of converting global warming into one or more of the other categories further simplifies the results. However, going beyond this and aggregating based on an arbitrary or subjective weighing which is hidden from the decision-maker does not serve to improve or simplify the results which will need to be reverted to their component parts if utilised by other LCA practitioners or decision-makers who may choose to apply their own weighing based on their own priorities.

Despite recommendations from the authors of widely used LCA methods, weighting was often used when it was unnecessary to do so. There were several examples of methods being insufficiently explained, unweighted data being concealed preventing third parties from making their own comparison, and other examples of bad practice. None of the 13 studies which used a single unit score for comparison gave a rationale for their use of normalisation or weighting and most omitted the weightings used.

Several recommendations concerning the utilisation of single impact categories or aggregation are made in the next section, informed by the results of the review.

5. Recommendations

5.1. Methodological Considerations

Single impact category approach

Where time and resources are limited, it may be preferable to measure a single impact category, such as energy demand, for short-term decisions. In the comparisons from the literature review, this indicator had a direct relationship with the most frequently used indicators (global warming potential, acidification potential and eutrophication potential) in more than 93% of comparisons. This method should be used with caution, however, as the unmeasured cost could far outweigh the apparent saving in a single impact category (e.g. Saleh, 2016). For long-term, large-scale projects the inclusion of additional categories is recommended.

Weighting and normalisation of damage categories

Many of the studies claiming to use the widely used LCA methods have ignored vital prescriptions regarding weighting and normalisation. It is therefore necessary to highlight these for future studies including LCAs. It is recommended, as previously stated in the CML method (Guinée et al. 2002) but ignored in many of the studies in this review, that weighting is not used where it can be avoided. Where it is applied, the normalisation and weighting factors or the pre-normalisation data must be provided so the objective results can be discerned. Furthermore, endpoint or damage categories should only be used where a definite preference cannot be determined using only midpoint categories or inventory inputs.

5.2. Further Research

A "benefit" indicator approach

Upon reviewing previous LCA papers, it is apparent that there is an appetite to develop a universal indicator; a single unit of impact with which to measure and compare the performance of a given set of products or processes. Guinée et al., (2002) recommend further examination of impact category weighting and argue for a more definite set of weighting factors to standardise the LCA method. To avoid subjectivity in the weighting of different categories, and the paradoxical effect of unweighted normalised endpoint categories, it is suggested that the use of normalised end point benefit rather than end point damage categories (e.g. life expectancy, biodiversity). With this approach, an endpoint category would have a greater impact on an LCA comparison the more it is threatened by background damage.

Objective approach to LCA weighting method.

It might be the case that a method that extrapolates all midpoint indicators to human survival (or one of the other endpoint categories) using objective scientific data is more appropriate than attempting to weight distinct categories. In such as method, the only value judgment would be that of timescale which would not necessarily be made at the point of presenting results, but could be made by decision-makers in possession of the objective findings. Sources that attempt to convey the seriousness of the environmental crisis frequently make this extrapolation in order the illustrate their point (e.g. Wallace-Wells, 2019).

6. Declarations

Funding and/or Conflicts of interests/Competing interests

The authors declare no conflict of interest. No funding has been received in support of this paper.

7. References

Arvidsson, R. and Svanstrom, M., 2015. ‘A Framework for Energy Use Indicators and Their Reporting in Life Cycle Assessment’, Integrated Environmental Assessment and Management — Volume 12(3), pp.429–436. https://doi.org/10.1002/ieam.1735

Banar M. and Cokaygil Z., 2008. ‘A comparative life cycle analysis of two different juice packages’, Environmental Engineering Science Volume 25(4), pp.549-555. https://doi.org/10.1089/EES.2007.0079

Bertolini, M., Bottani, E., Vignali, G. and Volpi, A., 2016. ‘Comparative Life Cycle Assessment of Packaging Systems for Extended Shelf Life Milk’, Packaging Technology and Science, 2016(29), pp.525–546. https://doi.org/10.1002/pts.2235

Bertoluci, G., Leroy, Y. and Olsson, A., 2014. ‘Exploring the environmental impacts of olive packaging solutions for the European food market’, Journal of Cleaner Production 64, pp.234-43. https://doi.org/10.1016/j.jclepro.2013.09.029

Björklund, A. E., 2002. ‘Survey of Approaches to Improve Reliability in LCA’, The International Journal of Life Cycle Assessment 7(2), pp.64-72. https://doi.org/10.1007/BF02978849

Bohlmann, G.M., 2004. ‘Biodegradable Packaging Life-Cycle Assessment’, Environmental Progress Vol. 23(4), pp.342-346. https://doi.org/10.1002/ep.10053

British Standards Organisation, 2018. ‘BS EN ISO 14044:2006+A1:2018: Environmental management - Life cycle assessment - Requirements and guidelines’, London: Publisher.

Bugnicourt E., Schmid M., Nerney O.Mc., Wildner J., Smykala L., Lazzeri A. and Cinelli P., 2013. ‘Processing and validation of whey-protein-coated films and laminates at semi-industrial scale as novel recyclable food packaging materials with excellent barrier properties’, Advances in Materials Science and Engineering Volume 2013, Article ID 496207. https://doi.org/10.1155/2013/496207

Cabot M.I., Luque A., De Las Heras A. and Aguayo F., 2019. ‘Aspects of sustainability and design engineering for the production of interconnected smart food packaging’ PLoS ONE 14(5): e0216555. https://doi.org/10.1371/journal.pone.0216555

Cinelli, P., Coltelli, M. B., Mallegni, N., Morganti, P., and Lazzeri, A., 2017. ‘Degradability and Sustainability of Nanocomposites Based on Polylactic Acid and Chitin Nano Fibrils’, Chemical Engineering Transactions Vol.60(2017). https://doi.org/10.3303/CET1760020

De Monte M., Padoano E., Pozzetto D., 2005. ‘Alternative coffee packaging: An analysis from a life cycle point of view’, Journal of Food Engineering 66(2005), pp.405–411. https://doi.org/10.1016/j.jfoodeng.2004.04.006

De Monte, M., Padoano, E. and Pozzetto, D., 2005. ‘Alternative coffee packaging: an analysis from a life cycle point of view’, Journal of Food Engineering 66(2005), pp.405–411. https://doi.org/10.1016/j.jfoodeng.2004.04.006

De Schryver, A.M., Brakkee, K.W., Goedkoop, M.J. and Huijbregts M.A.J. 'Characterization Factors for Global Warming in Life Cycle Assessment Based on Damages to Humans and Ecosystems', Environmental Science & Technology vol.43(6), pp.1689-1695. https://doi.org/10.1021/es800456m

Finnveden, G., 1997. 'Valuation Methods Within LCA - Where are the Values?' The International Journal of Life Cycle Assessment 2(3), pp.163-169. https://doi.org/10.1007/BF02978812

Fresán, U., Errendal, S., Craig, W.J. and Sabaté, J., 2019. 'Does The Size Matter? A Comparative Analysis Of The Environmental Impact Of Several Packaged Foods', Science of the Total Environment 687, pp.369-379. https://doi.org/10.1016/j.scitotenv.2019.06.109

Frischknecht, R., Wyss, F., Knöpfel, S.B., Lützkendorf, T., Balouktsi, M., 2015. 'Cumulative energy demand in LCA: the energy harvested approach', The International Journal of Life Cycle Assessment 2015(20), pp.957–969. https://doi.org/10.1007/s11367-015-0897-4

Fruergaard, T., Ekvall, T. and Astrup, T., 2009. 'Energy use and recovery in waste management and implications for accounting of greenhouse gases and global warming contributions', Waste Management & Research 2009(27), pp.724–737. https://doi.org/10.1177/0734242X09345276

Geyer et al., R., Jambeck, J.R. and Law, K.L., 2017. 'Production, Use, and Fate of All Plastics Ever Made', Science Advances Vol.3(7). https://doi.org/10.1126/sciadv.1700782

Goedkoop, M., and Spriensma, R., 1999. 'The Eco-indicator 99: A damage oriented method for Life Cycle Impact Assessment Methodology Report.', Amersfoort: PRé Consultants.

Goedkoop, M., and Spriensma, R., 2001. 'The Eco-Indicator'99. A Damage Oriented Method for Life Cycle Impact Assessment, Methodology Report. 3rd Edition, PRé Consultants, Amersfoort.

Goedkoop, M., Heijungs, R., Huijbregts, M., De Schryver, A.M., Struijs, J. and van Zelm, R., 2009. 'ReCiPe 2008: A life cycle impact assessment method which comprises harmonised category indicators at the midpoint and the endpoint level First edition' Amersfoort: PRé Consultants.

Guinée J.B., Gorrée M., Heijungs R., Huppes G., Kleijn R., van Oers L., Wegener Sleeswijk A., Suh S., Udo de Haes H.A., de Bruijn H., van Duin R., Huijbregts M.A.J., 2002. 'Life Cycle Assessment: An Operational Guide to the ISO Standards'. Dordrecht: Kluwer Academic Publishers.

Guiso, A., Parenti, A., Masella, P., Guerrini, L., Baldi, F. and, Spugnoli, P., 2016. 'Environmental impact assessment of three packages for high-quality extra-virgin olive oil', Journal of Agricultural Engineering 2016; XLVII:515, pp.191-196. https://doi.org/10.4081/jae.2016.515

Huijbregts, M.A.J., Rombouts, L.J.A., Hellweg, S., Frischknecht, R., Jan Hendriks, A., van de Meent, D., Ragas, A.D.J., Reijnders, L. and Struijs, J., 2006. 'Is Cumulative Fossil Energy Demand a Useful Indicator for the Environmental Performance of Products.?', Environmental Science & Technology, 40(3), pp.641-648. https://doi.org/10.1021/es051689g

Huijbregts M.A.J., Hellweg, S., Frischknecht, R., Hendriks, H.W.M., Hungerbühler, K. and Jan Hendriks, A., 2010. 'Cumulative energy demand as predictor for the environmental burden of commodity production', Environmental Science & Technology, 2010, 44(6), pp.2189–2196. https://doi.org/10.1021/es902870s

Huijbregts M.A.J., Steinmann, Z.J.N., Elshout, P.M.F, Stam, G., Verones, F., Vieira, M.D.M., Hollander, A., Zijp, M. and van Zelm, R., 2016. 'ReCiPe 2016 v1.1: A harmonized life cycle impact assessment method at midpoint and endpoint level' Bilthoven: National Institute for Public Health and the Environment.

Humbert, S., De Schryver, A.M., Bengoa, X., Margni, M. and Jolliet, O., 2012. 'IMPACT 2002+: User Guide for version Q2.21', Lausanne: Quantis.

Ingrao, C., Lo Giudice, A., Bacenetti, J., Khaneghah, A.M., Sand'Ana, A.S., Rana, R. and Siracusa, V., 2015. 'Foamy polystyrene trays for fresh-meat packaging: Life-cycle inventory data collection and environmental impact assessment', Food Research International 76, pp.418-426. https://doi.org/10.1016/j.foodres.2015.07.028

Ingrao, C., Gigli, M. and Siracusa, V., 2017. ‘An attributional Life Cycle Assessment application experience to highlight environmental hotspots in the production of foamy polylactic acid trays for fresh-food packaging usage’, Journal of Cleaner Production 150, pp.93-103. https://doi.org/10.1016/j.jclepro.2017.03.007

Ingarao, G., Licata, S. Sciortino, M., Planeta, D., Di Lorenzo, R. and Fratini, L., 2017. ‘Life cycle energy and CO2 emissions analysis of food packaging: an insight into the methodology from an Italian perspective’, International Journal of Sustainable Engineering Vol.10(1), pp.31-43. https://doi.org/10.1080/19397038.2016.1233296

Jensen, A.A., Hoffman, L., Møller, B.T. and Schmidt, A., 1997. ‘Life Cycle Assessment: A guide to approaches, experiences and information sources’ Environmental Issues Series(6), Copenhagen: European Environment Agency.

Joachimiak-Lechman, K., Selech, J. and Kasprzak, J., 2018. ‘Eco-efficiency analysis of an innovative packaging production: case study’, Clean Technologies and Environmental Policy 2019(21), pp.339–350. https://doi.org/10.1007/s10098-018-1639-7

Jolliett, O., Margni, M., Charles, R., Humbert, S., Payet, J., Rebitzer, G. and Rosenbaum, R., 2003. ‘Impact 2002+: A New Life Cycle Impact Assessment Methodology’, The International Journal of Life Cycle Assessment 8, pp.324-330. http://doi.org/10.1007/BF02978505

Miettinen, P. and Hämäläinen, R. P., 1997. ‘How to benefit from decision analysis in environmental life cycle assessment (LCA)’, European Journal of Operational Research 102(1997), pp.279-294. https://doi.org/10.1016/S0377-2217(97)00109-4

Nalbone, L., Cincotta, F., Giarratana, F. and Panebianco, G.Z.A., 2021. ‘Microplastics in Fresh and Processed Mussels Sampled From Fish Shops and Large Retail Chains in Italy’, Food Control 125, 108003. https://doi.org/10.1016/j.foodcont.2021.108003

Puig, R., Fullana-i-Palmer, P., Baquero, G., Riba, J. and Bala, A., 2013. ‘A Cumulative Energy Demand indicator (CED), life cycle based, for industrial waste management decision making’, Waste Management 33, pp.2789-2797. https://doi.org/10.1016/j.wasman.2013.08.004

Rahimi, A. and Garcia, J.M., 2017. ‘Chemical recycling of waste plastics for new materials production’, Nature Reviews Chemistry Vol.1, Article 0046. https://doi.org/10.1038/s41570-017-0046

Sala, S. and Andreasson, J., 2018. ‘Improving Interpretation, Presentation and Visualisation of LCA Studies for Decision Making Support’ in Benetto, E., Gericke, K. and Guiton, M. (eds), Designing Sustainable Technologies, Products and Policies. Cham: Springer, pp.337-342. https://doi.org/10.1007/978-3-319-66981-6

Saleh, Y., 2016. ‘Comparative life cycle assessment of beverages packages in Palestine’, Journal of Cleaner Production 131(2016), pp.28-42. https://doi.org/10.1016/j.jclepro.2016.05.080

Siracusa, V., Ingrao, C., Lo Giudice, A., Mbohwa, C. and Marco Dalla Rosa, M., 2014. ‘Environmental assessment of a multilayer polymer bag for food packaging and preservation: An LCA approach’, Food Research International 62(2014), pp.151-161. http://doi.org/10.1016%2Fj.foodres.2014.02.010

Vidal, R., Martínez, P., Mulet, E., González, R., López-Mesa, B., Fowler, P. and Fang, J.M., 2007. ‘Environmental assessment of biodegradable multilayer film derived from carbohydrate polymers’, Journal of Polymers and the Environment, 2007(15), pp.159–168. http://doi.org/10.1007/s10924-007-0056-5

Wallace-Wells, D., 2019. ‘The Uninhabitable Earth: Life After Warming’, New York: Tim Duggan Books.

Wang, J., Liu, X., Li, Y., Powell, T., Wang, X., Wang, G., Zhang, P., 2019. ‘Microplastics as contaminants in the soil environment: A mini-review’, Science of the Total Environment 691, pp.848-857. https://doi.org/10.1016/j.scitotenv.2019.07.209

Wikström, F., Williams, H., Verghese, K. and Clune, S., 2014. ‘The influence of packaging attributes on consumer behaviour in food-packaging life cycle assessment studies - a neglected topic’, Journal of Cleaner Production 73(2014), pp.100-108. http://doi.org/10.1016/j.jclepro.2013.10.042

Zhang, H., Hortal, M., Dobon, A., Bermudez, J.M. and Lara-Lledo, M., 2015. ‘The Effect of Active Packaging on Minimizing Food Losses: Life Cycle Assessment (LCA) of Essential Oil Component-enabled Packaging for Fresh Beef’, Packaging Technology and Science, 2015(28), pp.761–774. https://doi.org/10.1002/PTS.2135

7. Tables

Table 1: Various impact categories used in LCA studies

Impact category	Abbreviation	Unit	Inclusions in studies	Total comparisons	Total comparisons which also included En
Global warming potential, climate change, GHG emissions	GWP	CO_2 eq.	29	278	107
Acidification potential, acidification (terrestrial, aquatic)	AP	SO_2 eq., mol H+ eq., species/year	20	144	88
Photochemical oxidant creation potential, photochemical ozone formation	POC	C_2H_4 eq., NMVOC eq., ethane-eq., NOx eq.	15	171	75
Ozone depletion	ODP	R11 eq., CFC-11 eq., DALY	14	113	74
Eutrophication potential	EP	PO_4 eq., mol N eq.	12	115	81
Primary Energy Use, Cumulative energy demand, energy use (renewable, non-renewable), resource consumption	En	MJ	12	107	NA
Freshwater ecotoxicity	FWE	1,4-DCB eq., species/year, CTUe	8	95	3
Human toxicity potential	HTP	1,4-DB eq., C_2H_4 eq., DALY, DCB eq.	7	77	9
Particulate matter	PM	$PM_{2.5}$ eq., DALY, Disease incidence	7	21	7
Fossil fuels	Fos	oil eq., $	6	72	1
Terrestrial ecotoxicity	TE	1,4-DCB eq., species/year	6	28	3
Marine aquatic ecotoxicity	MaE	1,4-DCB eq., species/year	6	28	3
Water resource depletion, water consumption	WRD	m^3	6	14	9
Freshwater eutrophication	FEu	P eq., species/year	6	12	6
Abiotic Depletion Potential, Mineral, fossil & ren resource depletion	ADP	Sb eq.	4	25	3
Non-carcinogens	NCar	1,4-DCB eq., toluene eq.	4	21	9
Carcinogens	Car	1,4-DCB eq., benzene eq.	4	13	0
Mineral resource	Min	Cu eq.	4	12	9
Ionizing radiation	IR	kBq Co60 eq., DALY	4	5	0
Land use	LU	C deficit, m2a crop eq., m^2	4	4	0
Marine eutrophication	MEu	N eq	3	4	0
Human toxicity (soil)	HTs	m^3 soil, kg	2	13	3
Respiratory effects	Res	kg, $PM_{2.5}$ eq.	2	13	3
Ecotoxicity	ETX	2,4-D eq	2	11	0
Smog	Sm	NOx eq., O_3-eq.	2	11	0

Organic respiratory compounds	OrC	%	2	10	9
Human health	HH	PM_{10}-eq., DALY	2	7	6
Metal depletion	Met	Fe eq., $	2	3	0

Table 2: Trade-off (MJ/kg CO2) values in various studies

Comparison	**Δ_E Energy demand (MJ)**	**Δ_G GWP (kg CO2 eq.)**	**Trade-off Δ_E/Δ_G (MJ/kg CO2)**	**Trade-off acceptable by authors?**
Bohlmann, 2004	1036	3	**345.333(high)**	Not stated
Guiso et al., 2016_1	1.494	0.087	17.1724	Not stated
Guiso et al., 2016_2	1.526	0.049	31.1429	Not stated
Guiso et al., 2016_3	0.554	0.127	4.3622	Not stated
Saleh, 2016	2964.3	2113.58	**1.4025(low)**	Authors recommend lower energy demand. <1.4025
Zhang et al., 2015	0.284	0.005	56.8	Not stated
		Mean	76.0	
		σ	133.5	

Table 3: Trade-off (MJ/g SO2) values in various studies

Comparison	**Δ_E Energy demand (MJ)**	**Δ_{AP} AP (g SO2 eq.)**	**Trade-off Δ_E/Δ_{AP} (MJ/g SO2)**	**Trade-off acceptable by authors?**
Guiso et al., 2016	0.019	0.388	**0.049(low)**	Not stated
Vidal et al., 2007	4.5	1.58	**2.848(high)**	Authors recommend lower energy demand. <2.8481
Zhang et al., 2015	0.284	0.2	1.42	Not stated
		Mean	1.439	
		σ	1.14	

Table 4: Trade-off (MJ/g PO4) values in various studies

Comparison	**Δ_E Energy demand (MJ)**	**Δ_{EP} EP (g PO4 eq.)**	**Trade-off Δ_E/Δ_{EP} (MJ/g PO4)**	**Trade-off acceptable by authors?**
Vidal et al., 2007	4.5	0.391	11.509	Authors recommend lower energy demand. <11.509
Zhang et al., 2015	0.284	0.1	2.84	Not stated

Bertolini et al., 2016	0.7	0.0118	**59.322(high)**	Not stated
Guiso et al., 2016_1	0.051	1.494	**0.034(low)**	Not stated
Guiso et al., 2016_2	0.145	0.554	0.261	Not stated
		Mean	14.793	
		σ	25.325	